

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
13. Oktober 2005 (13.10.2005)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2005/095446 A1

(51) Internationale Patentklassifikation⁷: **C07K 14/195**

AUF AKTIEN [DE/DE]; Henkelstr. 67, 40589 Düsseldorf (DE).

(21) Internationales Aktenzeichen: PCT/EP2005/001543

(22) Internationales Anmeldedatum:
16. Februar 2005 (16.02.2005)

(72) Erfinder; und

(25) Einreichungssprache: Deutsch

(75) Erfinder/Anmelder (nur für US): **FEESCHE, Jörg** [DE/DE]; Georg-Büchner Str. 23, 40699 Erkrath (DE). **MEINHARDT, Friedhelm** [DE/DE]; Erlengrund 207, 48308 Senden (DE). **NAHRSTEDT, Hannes** [DE/DE]; Martin-Niemöller-Str. 1, 48159 Münster (DE). **WALDECK, Jens** [DE/DE]; Im Wieshof 8, 58708 Menden (DE). **GROENE, Mark** [DE/DE]; Berlinger Str. 33, (App. 1110), 55101 Mainz (DE). **EICHSTÄDT, Renee** [DE/DE]; Auerstrasse 11, 50733 Köln (DE).

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
10 2004 013 988.1 19. März 2004 (19.03.2004) DE

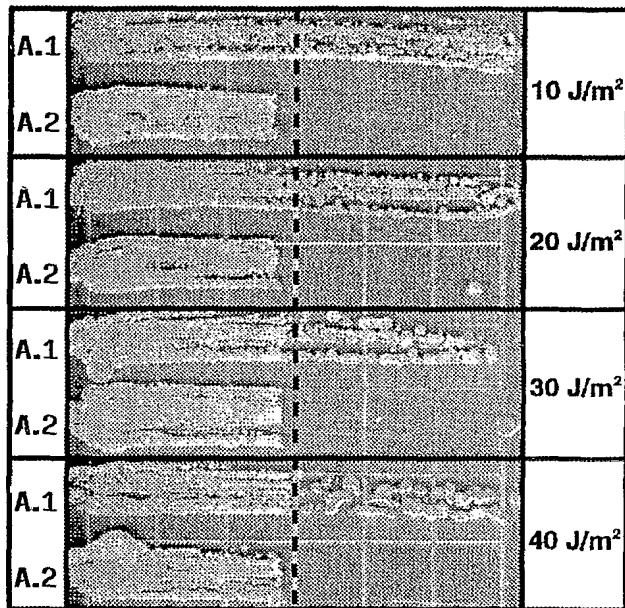
(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): **HENKEL KOMMANDITGESELLSCHAFT**

[Fortsetzung auf der nächsten Seite]

(54) Title: FACTOR RECA FROM BACILLUS LICHENIFORMIS AND RECA-INACTIVATED SAFETY STEMS USED FOR BIOTECHNOLOGICAL PRODUCTION

(54) Bezeichnung: DER FAKTOR RECA AUS BACILLUS LICHENIFORMIS UND RECA-INAKTIVIERTE SICHERHEITSTÄMME FÜR DIE BIOTECHNOLOGISCHE PRODUKTION

B



(57) Abstract: The invention relates to the factor RecA from *Bacillus licheniformis* DSM 13 (SEQ ID NO. 2), along with the associated gene *recA* (SEQ ID NO. 1), including related proteins and genes thereof, such as the variant indicated under SEQ ID NO. 31 and 32, among others. According to the invention, gene *recA* is used for constructing gram-positive bacterial safety stems for biotechnological production, among other things, by inactivating the same in the respective stems. In a special embodiment, said stems are provided with additional functional deletions in phase-IV sporulation genes, preferably in gene *spoIV* (in *Bacillus licheniformis*), gene *yqfD* (in *B. subtilis*), or the respective gene that is homologous thereto if said stems are naturally able to form spores. Furthermore, the inventive RecA represents a protein which can be used in molecular biological assays or for modulating the molecular biological activities of cells, especially in connection with DNA polymerization or recombination processes.

aus *Bacillus licheniformis* DSM 13 (SEQ ID NO. 2), zusammen mit dem zugehörigen Gen *recA* (SEQ ID NO. 1), einschließlich verwandten Proteinen und Genen hierzu, darunter der unter SEQ ID NO. 31 beziehungsweise 32 angegebenen Variante. Das Gen *recA* wird erfindungsgemäß unter anderem zur Konstruktion von grampositiven bakteriellen Sicherheitsstämmen für die biotechnologische Produktion eingesetzt, indem es in den betreffenden

(57) Zusammenfassung: Gegenstand der vorliegenden Anmeldung ist der Faktor RecA

[Fortsetzung auf der nächsten Seite]

WO 2005/095446 A1



(81) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG,

ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— mit internationalem Recherchenbericht

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

Der Faktor RecA aus *Bacillus licheniformis* und recA-inaktivierte Sicherheitsstämme für die biotechnologische Produktion

Die vorliegende Erfindung betrifft den Faktor RecA aus *Bacillus licheniformis* sowie Mikroorganismen als Sicherheitsstämme für die biotechnologische Produktion, die dadurch gekennzeichnet sind, daß sie funktionelle Deletionen in dem zugehörigen Gen *recA* aufweisen. Ferner steht RecA damit für weitere molekularbiologische Ansätze zur Verfügung.

Die vorliegende Erfindung liegt auf dem Gebiet der Biotechnologie, insbesondere der Herstellung von Wertstoffen durch Fermentation von Mikroorganismen, die zur Bildung der interessierenden Wertstoffe in der Lage sind. Hierzu zählen beispielsweise die Herstellung niedermolekularer Verbindungen, etwa von Nahrungsmittelergänzungsstoffen oder pharmazeutisch relevanten Verbindungen, oder von Proteinen, für welche aufgrund ihrer Diversität wiederum ein großes technisches Einsatzgebiet besteht. Im ersten Fall werden die Stoffwechseleigenschaften der betreffenden Mikroorganismen zur Herstellung der Wertstoffe ausgenutzt und/oder verändert; im zweiten Fall werden Zellen eingesetzt, die die Gene der interessierenden Proteine exprimieren. In beiden Fällen handelt es sich zumeist also um gentechnisch veränderte Organismen (GVO).

Zur Fermentation von Mikroorganismen besteht ein reichhaltiger Stand der Technik, insbesondere auch im großtechnischen Maßstab; er reicht von der Optimierung der betreffenden Stämme hinsichtlich der Bildungsrate und der Nährstoffausnutzung über die technische Gestaltung der Fermenter bis hin zur Gewinnung der Wertstoffe aus den betreffenden Zellen selbst und/oder dem Fermentationsmedium. Hierfür kommen sowohl genetische und mikrobiologische als auch verfahrenstechnische und biochemische Ansätze zu tragen. Ziel der vorliegenden Erfindung ist es, diesen Prozeß hinsichtlich der sicherheitsrelevanten Aspekte der eingesetzten Mikroorganismen zu verbessern, und zwar auf der Ebene der genetischen Eigenschaften der betrachteten Stämme.

Dies steht vor dem Hintergrund, daß die Verwendung gentechnisch veränderter Organismen im allgemeinen strengen gesetzlichen Richtlinien bezüglich der biologischen Sicherheit unterliegt. In den meisten Ländern sind die Betreiber von Anlagen mit GVO

gehalten, dafür zu sorgen, daß nach Möglichkeit keine GVO in die Umgebung gelangen. Zusätzlich sollen für die Produktion verwendete GVO Eigenschaften aufweisen, die es ihnen – falls sie doch in die Umgebung gelangen sollten – erschweren oder je nach Gefahrenpotential sogar unmöglich machen sollen, sich dort zu vermehren („Containment“-Konzept).

Dem Übersichtsartikel „Suicidal genetic elements and their use in biological containment of bacteria“ von S.Molin et al. (*Annu. Rev. Microbiol.*, 1993, Band 47, Seiten 139 bis 166) zufolge wird dabei zwischen den beiden grundsätzlichen Strategien differenziert, als „aktive“ Komponenten kontrollierte Suizidsysteme in die Zellen einzuführen oder über „passive“ Systeme die Zelleigenschaften so zu verändern, daß ihre Überlebenschancen unter Stressbedingungen sinken. Die zweite, für die vorliegende Anmeldung relevante Strategie wird darin auch als „Disablement approach“ bezeichnet.

Stämme von GVO mit einem verminderten Risiko für Mensch und Umwelt im Falle einer unbeabsichtigten Freisetzung werden als Sicherheitsstämme bezeichnet. Je nach den grundsätzlichen Eigenschaften der Mikroorganismen werden zunehmend mehrere Eigenschaften gefordert, von denen jede für sich einen Sicherheitsaspekt darstellt. Somit ist es vorteilhaft, verschiedene Instrumente zur Erzeugung von Sicherheitsstämmen zur Verfügung zu haben. Hiervon sind einige „passive“ Systeme bereits im Stand der Technik beschrieben.

So betrifft die Anmeldung EP 369817 A1 *Bacillus*-Stämme, insbesondere *B. subtilis*, zur Herstellung und Sekretion von Proteinen, bei denen die Gene für extra- und intrazelluläre Proteasen, nämlich *epr*, *rp-I*, *rp-II*, *isp-1*, *apr* und/oder *npr* durch Punktmutationen oder Insertionen inaktiver Genkopien funktionell inaktiviert worden sind. Der Sinn dieser gentechnischen Veränderungen besteht darin, die für die mit diesen Stämmen hergestellten, interessierenden Proteine schädlichen Protease-Aktivitäten zu minimieren. Die betreffenden Stämme können zusätzlich über Mutationen verfügen, die die Sporulation und damit eine Bildung ebenso schädlicher Sporulationsproteasen verhindern. Hierunter wird das in Phase Null der Sporulation (siehe unten) von *B. subtilis* aktive Gen *spoOA* genannt, um durch dessen Inaktivierung die mit der Sporulation verbundene Bildung intrazellulärer Proteasen zu verhindern.

Die Anmeldung WO 92/16642 A1 verfolgt den gleichen Lösungsansatz: sie offenbart, daß durch Inaktivierung der Proteasegene *apr*, *npr*, *isp-1*, *epr*, *bpr*, *rsp* und *mpr* von *Bacillus* ein Großteil der extrazellulären Protease-Aktivität ausgeschaltet wird, und lehrt, daß dies durch die Inaktivierung des neu beschriebenen Gens *vpr* für die Rest-Protease III noch verbessert werden kann. Auch hier wird auf die Möglichkeit der Inaktivierung von *spoOA* hingewiesen, um die Bildung intrazellulärer Proteasen zu verhindern.

Bei der Sporulation von grampositiven Bakterien handelt es sich um einen Entwicklungsprozeß zur Bildung von Dauerformen, den sogenannten Sporen, zum Überdauern widriger Umwelteinflüsse. Sie wird über eine komplexe Regulationskaskade mit vermutlich mehr als 100 Genen und unter Beteiligung von speziellen Sigma-Faktoren gesteuert. Den Zusammenhang dieses Prozesses mit dem Zellzyklus von *B. subtilis* beschreibt beispielsweise die Publikation „Cell cycle and sporulation in *Bacillus subtilis*“ (1998) von P.A. Levin und A.D. Grossmann in *Curr. Opin. Microbiol.*, Band 1, Seiten 630 bis 635. Hier wird vor allem der Transkriptionsfaktor Spo0A als Kontrollelement zum Einleiten der Sporulation dargestellt. Die sequentielle Aktivierung der phasenspezifischen Gene durch verschiedene Sigmafaktoren faßt beispielsweise der Übersichtsartikel „Control of sigma factor activity during *Bacillus subtilis* sporulation“ (1999) von L. Kroos et al. in *Mol. Microbiol.*, Band 31, Seiten 1285 bis 1294 zusammen. Bei diesem Vorgang werden ihrer Reihenfolge nach die aufeinanderfolgenden Stadien Null und dann I bis VII beobachtet. Diese Numerierung findet sich auch in den Bezeichnungen der beteiligten Gene und Faktoren wieder.

Die Anmeldung EP 492274 A2 offenbart, daß im Stand der Technik bereits über unspezifische Mutagenese die Inaktivierung von Sporulationsgenen gelungen sei, wodurch asporogene Mutanten (spo-Minus-Phänotyp) erhalten worden seien. EP 492274 A2 selbst beschreibt einen durch gezielte Mutagenese in dem frühen Sporulationsgen *spoID* behandelten *B. subtilis*-Stamm, der mit einer Reversionsrate von weniger als 10^{-8} praktisch nicht mehr in der Lage ist, Sporen zu bilden. Diese Anmeldung lehrt, diesen Stamm erst nach Inaktivierung der weiteren Gene *leu* (für die Leucin-Synthese), *pyrD1* (für die Uracil-Synthese), *apr* und *npr* für die Herstellung von Wertstoffen für die biotechnologische Produktion zu verwenden, weil hiermit Vorteile in der Produktion sowie Sicherheitsaspekte verbunden seien.

Die Anmeldung WO 97/03185 A1 befaßt sich ebenfalls mit der Inaktivierung der Sporulationsfähigkeit von *Bacillus*-Spezies mit Ausnahme von *B. subtilis* und Verwendung dieser Stämme zur biotechnologischen Herstellung von Wertstoffen. Dieser Anmeldung zufolge soll das frühe, für den Sigmafaktor F codierende Gen *spoIIAC* funktionell inaktiviert werden, vorteilhafterweise in Kombination mit Deletionen in Genen der ebenfalls früh aktivierte Sporulationsgruppen *spo2*, *spo3*. Hierfür wird eine irreversible Inaktivierung der betreffenden chromosomalen Abschnitte für *spoIIAC* beschrieben.

Die Anmeldung WO 02/097064 A1 (EP 1391502 A1) betrifft Mikroorganismen, bei denen Gene aus den Stadien II, III, IV oder V der Sporulation deletiert oder inaktiviert worden sind. Hierbei handelt es sich um die Gene *sigE*, *sigF*, *spoIIIE*, *spoIIISB* und *sigG* von *B. subtilis*, die innerhalb des Genorts von *spoIVCB* bis *spoIIIC* von *B. subtilis* liegen. Dieser kann anhand der Datenbank SubtiList (zugänglich über <http://genolist.pasteur.fr/SubtiList/genome.cgi>) auf den Bereich der Positionen von ca. 2.642.000 kb bis ca. 2.700.000 kb des inzwischen bekannten Gesamtgenoms von *B. subtilis* eingegrenzt werden. Dieser Anmeldung hatte die Aufgabe zugrundegelegen, überflüssige oder schädliche Aktivitäten von *Bacillus*-Stämmen auszuschalten, um die biotechnologische Produktion zu verbessern. Durch die genannten Modifikationen der mittleren bis späten Sporulationsgene würde bei Einsatz der betreffenden Stämme für die biotechnologische Produktion die Sporenbildung unterdrückt; dies wirke sich vorteilhaft auf die Nährstoff- und Energieverwertung aus; gleichzeitig könne die Dauer der Fermentation erhöht werden, um darüber die Gesamtausbeute an interessierendem Wertstoff zu erhöhen.

Der Übergang grampositiver Bakterien in die Dauerform der Spore wird also durch ungünstige Umweltbedingungen ausgelöst. Genau dies dürfte auch dann geschehen, wenn Bakterien aus den optimalen Wachstumsbedingungen der Fermentation unbeabsichtigt aus der Anlage austreten und in die Umgebung gelangen sollten. Demgegenüber ist, wie soeben dargelegt, über die Unterbindung der Fähigkeit zur Sporulation zur Erzeugung sicherer GVO bislang erst wenig nachgedacht worden. Der einschlägige Stand der Technik scheint lediglich nahezulegen, daß die Sporulation grampositiver Bakterien (a) wegen der damit verbundenen Proteaseaktivitäten und/oder (b) zur Verlängerung der Fermentationsdauer frühzeitig und vollständig unterbunden werden sollte, um letztlich die Fermentationsausbeute der so erhaltenen asporogenen

Stämme zu steigern. Zur Verfolgung von Sicherheitsaspekten werden dagegen nur mehrere, zusätzlich einzuführende Mutationen offenbart.

Das für den in Prokaryonten beschriebenen Faktor RecA codierende Gen *recA* ist in der Molekularbiologie sehr bekannt, bislang insbesondere jedoch aus einem anderen Zusammenhang als dem der Herstellung von Sicherheitsstämmen. Dieser Faktor bindet spezifisch und kooperativ an einzelsträngige DNA und sorgt unter ATP-Hydrolyse für eine teilweise Entwindung doppelsträngiger DNA. Dieser Vorgang ermöglicht den genetischen Vorgang der Rekombination, das heißt den Strangtausch zwischen ähnlichen DNA-Molekülen. So ist es in der Molekularbiologie ein übliches Vorgehen, das *recA*-Gen, dadurch zu verwenden, daß es über ein entsprechendes genetisches Konstrukt mit einer defekten *recA*-Kopie inaktiviert und somit ein *recA*-Minus-Phänotyp erzeugt wird, welcher nicht mehr zur Rekombination in der Lage ist. Beispielsweise nach dem Patent US 4713337 werden durch Crossing-over erzeugte Deletionsmutanten durch anschließende Inaktivierung von *recA* genetisch stabilisiert.

So tauchen Hinweise auf *recA* in den verschiedensten molekularbiologischen Zusammenhängen auf. Beispielsweise DE 10011358 A1, welche sich mit L-förmigen Bakterienstämmen befaßt, erwähnt zusätzlich, neben zahlreichen anderen möglichen Modifizierungen sei es unter anderem möglich, auch *recA* zu mutieren, um eine verbesserte Transformation und Plasmidstabilität zu erreichen.

Eine biochemische Beschreibung des RecA aus *Escherichia coli* liefert beispielsweise die Publikation „C-terminal deletions of the *Escherichia coli* RecA“ von S.L.Lusetti et al. (2003; *J. Biol. Chem.*, Band 278, Heft 18, Seiten 16372-16380). Daraus geht hervor, daß insbesondere der C-Terminus dieses Moleküls die Einzelstrangbindung vermittelt und entsprechende Deletionsmutanten in diesem Bereich neben anderen biochemischen Eigenschaften eine erhöhte Mitomycin-Sensitivität aufweisen. Mitomycin ist dafür bekannt, daß es die DNA-Synthese beeinträchtigt und dadurch bakterizid wirkt. Der N-terminale Bereich ist dagegen stärker an der Bindung von DNA-Doppelsträngen beteiligt.

Der bereits zitierte Übersichtsartikel von S.Molin et al. verweist ebenfalls auf Arbeiten, bei denen eine *recA*-Minus-Mutation als Marker für den gramnegativen *Escherichia coli* verwendet wird. Es wird die Vermutung geäußert, diese Mutation allein könne schon ausreichen, um jegliches Umweltrisiko durch diesen Stamm auszuschließen. Anderer-

seits werden zwei Nachteile dieses Ansatzes diskutiert, nämlich zum einen sei es technisch schwierig, diese Mutanten herzustellen und zum anderen würden die betreffenden Stämme auch in ihrem erwünschten Einsatz („short-term competitive properties“) derart behindert sein, daß man andere die Lebensfähigkeit beschränkende und im betreffenden Artikel erwähnte Mutationen vorziehen würde. Auch in Kombination mit dem grundsätzlich anderen Ansatz zur Erzeugung von Sicherheitsstämmen, nämlich dem Einführen von Suizidsystemen, habe sich die Inaktivierung von RecA als nachteilhaft herausgestellt.

Die Publikation „Freisetzung gentechnisch veränderter Bakterien“ von Selbitschka et al. (2003; *Biologie in unserer Zeit*, Band 33, Heft 3, Seiten 162-175) beschreibt die Freisetzung von gramnegativen, mit einem Luciferase-Gen modifizierten Bakterien der Spezies *Sinorhizobium meliloti* in einem mehrjährigen Freilandversuch. Sie trugen zusätzlich eine Inaktivierung des *recA*-Gens, welche dazu geführt hat, daß Zellen dieser Klone unter natürlichen Bedingungen letztlich nicht überleben konnten.

K.-D. Wittchen hat im Zuge seiner bei der Westfälischen Wilhelms-Universität Münster eingereichten Dissertation (1995) mit dem Titel „Entwicklung eines Sicherheitsstammes von *Bacillus megaterium* DSM 319 und molekulargenetische Charakterisierung des Gens für die extrazelluläre neutrale Metalloprotease (*nprM*)“ einen Stamm des grampositiven *B. megaterium* erzeugt, der nach gezielter Gendisruption Deletionen in der im Titel genannten neutralen Metalloprotease, der der Isopropylmalat-Dehydrogenase, und eines nicht näher bezeichneten SpolV-Proteins enthält. Anschließend wurde an diesem Stamm eine *recA*-Mutation vorgenommen, welche in bezug auf die UV-Sensitivität des Stammes im Vergleich zum Wildtyp allerdings keine wesentlichen Unterschiede verursachte, wohl aber bei Wachstum auf Mitomycin C-haltigen Agarplatten. Diese Vierfachmutante wurde als Sicherheitsstamm vorgeschlagen, ohne jedoch dessen Realisierbarkeit oder sogar die konkreten Auswirkungen dieser Modifikationen auf einen Produktionsprozeß mit entsprechend modifizierten Bakterienstämmen zu untersuchen.

Die in derselben Arbeitsgruppe angefertigte Diplomarbeit von H. Nahrstedt (2000) mit dem Titel „Molekulargenetische Charakterisierung des *recA*-Gens von *Bacillus megaterium* DSM 319 und Konstruktion einer Deletionsmutante“ schlägt folgende, nebeneinander vorliegende vier Mutationen in zum Teil Gruppen von Genen vor: *recA*-

Minus, Protease-Minus, Leucin-Auxotrophie und Sporulations-Defizienz. Es wird diskutiert, eine *recA*-Defizienz als eine Sicherheitsmarke neben anderen in Produktionsstämme einzuführen, weil sich dadurch zum einen unerwünschte Rekombinationsprozesse unterdrücken ließen und zum anderen die betreffenden Mutanten eine erhöhte Sensitivität gegenüber DNA-schädigenden Agenzien aufweisen, das heißt in der Umwelt eine geringere Überlebenschance besitzen sollten. Auch dieser Vorschlag wurde nicht weiterverfolgt.

Den Stand der Technik zu *RecA* kann man dahingehend zusammenfassen, daß dieses Protein bisher überwiegend aus genetischen Zusammenhängen bekannt ist. Der Einsatz dieses Faktors beziehungsweise die Inaktivierung des betreffenden Gens zur Erzeugung von Sicherheitsstämmen von GVO wird aufgrund seiner physiologischen Bedeutung bislang eher abgelehnt. Erfolgreiche beschriebene Beispiele hierfür sind lediglich *recA*-Minus-Mutanten der gramnegativen Spezies *Sinorhizobium meliloti* und des grampositiven *Bacillus megaterium*, letztere jeweils nur in Kombination mit drei weiteren sicherheitsrelevanten Mutationen.

Insgesamt bleibt festzuhalten, daß zur Herstellung von Sicherheitsstämmen für die biotechnologische Produktion verschiedene alternative genetische Systeme nach einem „passiven“ Wirkmechanismus etabliert sind, wie die Inaktivierung von Proteasegenen, die Ausschaltung von verschiedenen Stoffwechselgenen zur Erzeugung von Aminosäure- oder Nukleobasenauxotrophien. Bei sporenbildenden grampositiven Bakterien ist die Unterbindung der Sporulation beschrieben, insbesondere zu einem frühen Zeitpunkt, vor allem aber um zusätzliche Vorteile in der Fermentation zu erzielen. Dabei wird es als vorteilhaft angesehen, über mehrere, unterschiedlich wirkende Systeme zu verfügen, um sie nebeneinander auf einen bestimmten Stamm anzuenden, damit dieser als besonders sicher eingestuft werden kann.

Es stellte sich somit die Aufgabe, ein weiteres geeignetes Sicherheitssystem für gentechnisch veränderte grampositive Bakterien zu entwickeln, wofür als Grundlage zunächst ein geeigneter Faktor und/oder ein geeignetes Gen zu identifizieren waren.

Einen Teilaspekt dieser Aufgabe stellte nach Feststellung der grundsätzlichen Eignung eines solchen Systems die Isolierung eines hierfür verwendbaren genetischen Elements, eventuell eines Gens, und der Aminosäuresequenz eines hiervon gegebenenfalls

codierten Faktors dar, um dieses System entsprechenden molekularbiologischen Konstruktionen für den Einsatz in Produktionsstämmen zugänglich zu machen, insbesondere in Kombination mit einem oder mehreren weiteren der Sicherheit dienenden Regulationsmechanismen.

Ein weiterer Teilaспект der Aufgabe bestand darin, daß dieses System mit anderen Sicherheitssystemen kombinierbar sein sollte.

Somit lag eine Teilaufgabe darin, ein weiteres derartiges, hiermit zu kombinierendes Sicherheitssystem zu definieren, vorzugsweise eines, das neben diesen beiden Systemen keine weiteren Mutationen erforderlich machen würde. Mit anderen Worten: maximal diese zwei Mutationen sollten ausreichen, um einen grampositiven Sicherheitsstamm zu erzeugen, der weitgehende Anforderungen an die Verminderung der Lebensfähigkeit in der Umwelt erfüllen, das heißt zu einer minimalen Reversionsrate führen sollte. Denn eine niedrigere Zahl als vier nebeneinander wirksame Systeme bedeutet einen zunehmend geringeren Arbeitsaufwand zur Herstellung dieser Stämme.

Ein Nebenaspekt dieser Aufgabe bestand darin, ein derartiges Sicherheitssystem zu finden, das nicht so speziell ist, daß es nicht auch in anderen molekularbiologischen Ansätzen Verwendung finden könnte.

Diese Aufgabe wird durch den Faktor RecA mit einer Aminosäuresequenz, die zu der in SEQ ID NO. 2 angegebenen Aminosäuresequenz mindestens zu 96% identisch ist, beziehungsweise durch die für einen Faktor RecA codierende Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist, gelöst.

Die in SEQ ID NO. 2 und 1 angegebenen Aminosäure- und Nukleotidsequenzen sind die für RecA. Dabei codieren alle Positionen von 1 bis 1047 für das Protein; die letzten drei stellen dabei das Stop-Codon dar. Sie werden als Gen und Protein *recA* beziehungsweise RecA bezeichnet. Sie stammen aus dem bei der Deutschen Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig (<http://www.dsmz.de>) unter der Nummer DSM 13 hinterlegten Stamm *Bacillus licheniformis*. Erfindungsgemäße Lösungen der Aufgabe stellen alle Faktoren

beziehungsweise Nukleinsäuren dar, die hierzu eine hinreichend Homologie aufweisen, wie sie mit den jeweiligen Prozentangaben definiert ist.

Als nächster Stand der Technik kann der entsprechende Faktor aus *B. amyloliquefaciens* angesehen werden. Die zugehörigen vollständigen DNA- und Aminosäuresequenzen sind in der Datenbank NCBI der National Institutes of Health der USA (<http://www.ncbi.nlm.nih.gov>) unter der Zugangsnummer AJ515542 veröffentlicht, wobei eine nahe dem C-Terminus gelegene Teilsequenz zusätzlich aus dem Eintrag AY147924 hervorgeht. RecA aus *B. licheniformis* DSM 13 weist zum vollständigen Faktor auf Aminosäureebene eine Homologie von 94,0% Identität und auf Nukleinsäureebene eine Übereinstimmung von 81,2% Identität auf. Beide Vergleiche gehen aus den Alignments der Figuren 1 und 2 hervor, wo die Sequenzen von *B. amyloliquefaciens* jeweils in der zweiten Zeile dargestellt sind.

Als nächstähnliche Enzyme wurden RecA aus *B. subtilis* und RecE aus *B. subtilis* mit jeweils 93,4% Identität ermittelt. Sie weisen auf DNA-Ebene Homologiewerte von 81,0% beziehungsweise 81,2% Identität auf. Sie sind ebenfalls in der NCBI-Datenbank, und zwar unter den Eintragsnummern Z99112 (Region 161035 bis 162078) beziehungsweise X52132 veröffentlicht. Die Aminosäure- und DNA-Vergleiche mit diesen Faktoren sind ebenfalls in den Figuren 1 und 2 (jeweils Zeilen 3 beziehungsweise 4) dargestellt.

Wie weitere Faktoren RecA geeigneterweise erhalten werden können, die innerhalb des hier bezeichneten Homologiebereichs liegen, wird durch Beispiel 1 der vorliegenden Anmeldung illustriert. Dort ist ein molekularbiologisches Vorgehen gezeigt, wonach mithilfe bestimmter PCR-Primer, insbesondere den dort konkret offenbarten (SEQ ID NO. 25 bis 30) Oligonukleotiden, betreffende Gene beziehungsweise Genabschnitte aus chromosomal DNA-Präparationen der betreffenden Spezies gewonnen werden können. Gegebenenfalls können, sofern diese Primer nicht erfolgreich eingesetzt werden können, ähnliche Primer eingesetzt werden, bei denen – gesteuert über die Reaktionsbedingungen bei der Primer-Synthese – einzelne Positionen variiert werden. Diese PCR-Produkte lassen sich – falls bei Einsatz entsprechender Primer (vergleiche Figur 3B) nur Teilsequenzen erhalten worden sind – nach üblichen Methoden (Ausnutzung von Überlappungen) zu zusammenhängenden DNA-Sequenzen zusammensetzen. Hieraus ergibt sich unmittelbar die Aminosäuresequenz des von dem

erhaltenen Gens *recA* codierten Faktors RecA. Alternativ hierzu können auch die in SEQ ID NO. 1 oder SEQ ID NO. 31 offenbarten Sequenzen als Sonden verwendet werden, um nach an sich bekannten Methoden relevante Gene aus Genbanken zu isolieren.

Die hohen, beanspruchten Homologiewerte um den konkret hier beschriebenen Faktor lassen erwarten, daß derselbe Faktor RecA besonders in verwandten Stämmen oder Spezies, wahrscheinlich aber auch in weniger verwandten Spezies, wahrscheinlich sogar gramnegativen Organismen die zugehörige Funktion übernimmt. Diese liegt erfindungsgemäß in der eingangs erläuterten DNA-Einzelstrangbindung und der damit verbundenen Rolle für Rekombinationsvorgänge von Nukleinsäuren. Vergleichbare Wirkungen sollten auch mit Deletionen des *recA*-Gens verbunden sein, nämlich die Unterbindung von DNA-Rekombinationen und eine dadurch verminderte Lebensfähigkeit. Gleichzeitig sollte ein *recA*-Gen aus dem einen Stamm geeignet sein, diese Funktion in einem anderen zu übernehmen; dies wird mit zunehmender Ähnlichkeit zunehmend besser gelingen. Hierdurch wird der Einsatz des betreffenden Gens für die Herstellung von Deletionsmutanten der verschiedensten grampositiven Mikroorganismen möglich.

Dadurch wird für RecA und vor allem für das zugehörige Gen *recA* ein weites technologisches und kommerziell relevantes Gebiet eröffnet, nämlich die Herstellung von Wertstoffen durch Fermentation von genetisch veränderten grampositiven Bakterien. Sie können über Mutationen in *recA* nicht nur hinsichtlich ihrer genetischen Stabilität sondern auch ihrer Sicherheit verbessert werden. Dies gilt, wie unten weiter ausgeführt, insbesondere für Stämme, die in der biotechnologischen Produktion tatsächlich eingesetzt werden, wie beispielsweise *Bacillus licheniformis*.

Wie unten ebenfalls detaillierter ausgeführt geschieht dies vorzugsweise im Zusammenhang mit einer und besonders bevorzugt keinen weiteren sicherheitsrelevanten Deletionen. Ebenso bevorzugt geschieht dies in möglichst nahe verwandten Stämmen. Hierbei ist jedoch von *B. megaterium* abzusehen, zum einen weil hierfür, wie einleitend beschrieben, bereits die Spezies-eigenen *recA*-Gene beziehungsweise die hierin deletierten Mutanten zur Verfügung stehen und zum anderen weil diese Spezies, die sich insbesondere durch ihre großen Zellen und damit verbundene mikrobiologische Eigenheiten auszeichnet, im allgemeinen nicht für die großtechnische Fermentation genutzt wird.

Gegenstände der vorliegenden Erfindung liegen somit in dem Faktor RecA (SEQ ID NO. 2) und dem zugehörigen Gen *recA* (SEQ ID NO. 1) aus *B. licheniformis* DSM 13 beziehungsweise nahen Verwandten hierzu. Ebenso stellt die Verwendung eines solchen *recA*-Gens und/oder eines, das zu SEQ ID NO. 31 hinreichend verwandt ist, zur funktionellen Inaktivierung von *recA* in einem grampositiven Bakterium einen Erfindungsgegenstand dar, vorzugsweise in Kombination mit der funktionellen Inaktivierung eines in der Phase IV der Sporulation grampositiver Mikroorganismen aktiven Gens, vorzugsweise *spoIV*, *yqfD* beziehungsweise Homologen hierzu. Dies geschieht vorteilhafterweise mithilfe der in der vorliegenden Anmeldung zusätzlich beschriebenen Gene *spoIV* und *yqfD*. Einen entsprechenden Gegenstand der vorliegenden Erfindung stellen die hierdurch erhaltenen grampositiven Mikroorganismen dar; ebenso die mit diesen Organismen durchgeführten Fermentationen, insbesondere zur Herstellung von Wertstoffen. Des weiteren steht mit der vorliegenden Anmeldung ein RecA-Protein zur Verfügung, das in molekularbiologischen Ansätzen oder zur Modulation der molekularbiologischen Aktivitäten von Zellen verwendet werden kann, insbesondere im Zusammenhang mit DNA-Polymerisations- oder Rekombinationsvorgängen.

Zum ersten Erfindungsgegenstand gehört jeder oben definierte Faktor RecA mit einer Aminosäuresequenz, die zu der in SEQ ID NO. 2 angegebenen Aminosäuresequenz zunehmend bevorzugt mindestens zu 96,5%, 97%, 97,5%, 98%, 98,5%, 99%, 99,5% und ganz besonders bevorzugt zu 100% identisch ist.

Denn wie erläutert ist mit zunehmender Ähnlichkeit eine zunehmende Übereinstimmung der Funktionen und damit eine Austauschbarkeit der Faktoren zu erwarten.

Vorzugsweise handelt es sich dabei um einen Faktor RecA, der von einer Nukleinsäure codiert ist, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist.

In bevorzugten Ausführungsformen sind das Faktoren, die von einer Nukleinsäure codiert sind, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz zunehmend bevorzugt mindestens zu 87,5%, 90%, 92,5%, 95%, 96%, 97%, 98%, 99% und ganz besonders bevorzugt zu 100% identisch ist.

Denn über die Nukleinsäuren stehen die betreffenden Faktoren beziehungsweise Gene für die Transformation in andere, vorzugsweise verwandte Spezies oder für Modifikationen zur Verfügung. Hierzu gehören, wie unten detaillierter erläutert, insbesondere Mutationen der betreffenden Gene. Mit einem zunehmenden Maß an Identität zur angegebenen Sequenz sollte der Erfolg bei solchen Spezies umso größer sein, die zu *B. licheniformis* zunehmend verwandt sind, insbesondere bei der für die biotechnologische Produktion besonders wichtigen Spezies *B. licheniformis* selbst.

Die Gewinnung derartiger Nukleinsäuren geht wie oben erläutert aus Beispiel 1 hervor; auch auf die Isolierung aus Genbanken wurde bereits verwiesen.

Wie erläutert dienen der Verwirklichung der vorliegenden Erfindung vor allem die für einen Faktor RecA codierenden Nukleinsäuren, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist.

Dies gilt um so mehr für derartige Nukleinsäuren, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz zunehmend bevorzugt mindestens zu 87,5%, 90%, 92,5%, 95%, 96%, 97%, 98%, 99% und ganz besonders bevorzugt zu 100% identisch ist. Denn diese können über entsprechende Konstruktionen zur Transformation und/oder zur Mutagenese verwendet werden, wobei eine zunehmende Ähnlichkeit den erwünschten Erfolg umso wahrscheinlicher werden lässt.

Ganz besonders bevorzugt handelt es sich dabei um eine derartige Nukleinsäure, die für einen zuvor beschriebenen Faktor RecA codiert. Dies gilt beispielsweise für Strategien, bei denen ein funktioneller Faktor RecA hergestellt werden soll, etwa für die unten ausgeführten molekularbiologischen Versuchsansätze, oder für die Erzielung einer maximalen Übereinstimmung mit dem jeweils endogenen tatsächlich für RecA codierenden Gen, welches modifiziert und/oder ausgeschaltet werden soll. Denn in vielen Fällen reicht eine Mutation in einer einzigen Position, um etwa über eine nonsense-Mutation das Gen beziehungsweise den Faktor in seiner natürlichen Funktion auszuschalten.

Einen eigenen Erfindungsgegenstand stellt die Verwendung einer für einen Faktor RecA codierenden Nukleinsäure zur funktionellen Inaktivierung des Gens *recA* in einem grampositiven Bakterium dar, welches nicht *Bacillus megaterium* ist.

Denn zum einen gibt es für *B. megaterium* bereits die einleitend erwähnten Studien, in denen vorgeschlagen wird, gleichzeitig mit mehreren anderen Mutationen auch *recA* zu deletieren, um zu Sicherheitsstämmen zu gelangen. Zum zweiten stellen andere grampositive Bakterien wie beispielsweise solche der Gattungen *Bacillus*, *Staphylococcus*, *Corynebakterium* und *Clostridium* im Stand der Technik wichtigere Wirtsorganismen für die biotechnologische Produktion von Wertstoffen (siehe unten) dar.

Unter der funktionellen Inaktivierung ist im Sinne der vorliegenden Anmeldung jede Art von Modifikation oder Mutation zu verstehen, wonach die Funktion eines RecA als Einzelstrang-DNA-bindenden Faktors unterbunden wird. Dazu gehört die Ausführungsform, daß ein praktisch vollständiges, aber inaktives Protein gebildet wird, daß inaktive Teile eines RecA in der Zelle vorliegen, bis hin zu den Möglichkeiten, daß das Gen *recA* nicht mehr translatiert wird oder sogar vollständig deletiert ist. Somit besteht die eingangs diskutierte „Verwendung“ dieses Faktors oder dieses Gens dieser Ausführungsform nach darin, daß er beziehungsweise es von der betreffenden Zelle eben nicht mehr auf seine natürliche Weise zur Wirkung kommt. Dies wird diesem Erfindungsgegenstand zufolge auf genetischer Ebene dadurch erreicht, daß das betreffende Gen ausgeschaltet wird. Eine Möglichkeit, wie dies technisch erreicht werden kann, beschreibt Beispiel 2 der vorliegenden Anmeldung.

Nach einer Ausführungsform dieser Verwendung wird eine für ein nichtaktives Protein codierende Nukleinsäure mit einer Punktmutation eingesetzt.

Derartige Nukleinsäuren können über an sich bekannte Verfahren zur Punktmutagenese erzeugt werden. Solche sind beispielsweise in einschlägigen Handbüchern wie dem von Fritsch, Sambrook und Maniatis, „Molecular cloning: a laboratory manual“, *Cold Spring Harbour Laboratory Press*, New York, 1989, dargestellt. Zudem stehen hierfür inzwischen zahlreiche kommerzielle Baukästen zur Verfügung, etwa das QuickChange®-Kit der Firma Stratagene, La Jolla, USA. Das Prinzip besteht darin, daß Oligonukleotide mit einzelnen Austauschen (Mismatch-Primer) synthetisiert und mit dem einzelsträngig vorgelegten Gen hybridisiert werden; anschließende DNA-Polymerisation ergibt dann entsprechende Punktmutanten. Hierfür können die jeweiligen Spezies-eigenen *recA*-Sequenzen verwendet werden. Aufgrund der hohen Homologien ist es möglich und erfindungsgemäß besonders vorteilhaft, diese Reaktion anhand der mit SEQ ID NO. 1

zur Verfügung gestellten Sequenz oder etwa den anderen aus Figur 2 hervorgehenden Sequenzen verwandter Spezies durchzuführen. Diese Sequenzen können auch dazu dienen, entsprechende Mismatch-Primer für verwandte Spezies zu entwerfen, insbesondere anhand der im Alignment der Figur 2 identifizierbaren konservierten Bereiche.

Nach einer Ausführungsform dieser Verwendung wird eine Nukleinsäure mit einer Deletions- oder Insertionsmutation eingesetzt, vorzugsweise umfassend die jeweils mindestens 70 bis 150 Nukleinsäurepositionen umfassenden Randsequenzen des für das Protein codierenden Bereichs.

Auch diese Verfahren sind dem Fachmann an sich vertraut. Somit ist es möglich, die Bildung eines Faktors RecA durch die Wirtszelle dadurch zu verhindern, daß ein Teil des Gens auf einem entsprechenden Transformationsvektor über Restriktionsendonukleasen herausgeschnitten und der Vektor anschließend in den interessierenden Wirt transformiert wird, wo über die – bis dahin noch mögliche – homologe Rekombination das aktive Gen gegen die inaktive Kopie ausgetauscht wird. In der Ausführungsform der Insertionsmutation kann lediglich das intakte Gen unterbrechend oder anstelle eines *recA*-Genteils ein anderes Gen, beispielsweise ein Selektionsmarker eingefügt werden. Hierüber ist das Mutationereignis in an sich bekannter Weise genetisch und phänotypisch überprüfbar.

Solch ein Ansatz wurde in Beispiel 2 gewählt: Wie dort erläutert ist, wurden aus SEQ ID NO. 31 zwei flankierende Bereiche von jeweils ca. 340 bp ausgenutzt, um den dazwischenliegenden Teil des Gens *recA* eines *B. licheniformis*-Stamms zu deletieren (vergleiche Figur 3B). Im nachfolgenden Beispiel 3 wird der Erfolg dieser Deletion auf genetischer Ebene überprüft. So belegt Figur 4 (A und B), daß das betreffende DNA-Fragment durch die Deletion entsprechend verkürzt worden ist. Die phänotypische Beschreibung der dadurch erhaltenen Mutanten erfolgt in den nachfolgenden Beispielen. Demnach sind *recA*-inaktivierte Stämme deutlich UV-sensitiver als solche mit einem intakten *recA*-Gen (Beispiel 6).

Um diese jeweils notwendigen Rekombinationsereignisse zwischen dem in die Zelle eingeführten defekten Gen und der beispielsweise auf dem Chromosom endogen vorhandenen intakten Genkopie zu ermöglichen, ist nach dem derzeitigen Wissensstand eine Übereinstimmung in jeweils mindestens 70 bis 150 zusammenhängenden

Nukleinsäurepositionen, jeweils in den beiden Randsequenzen zu dem nichtübereinstimmenden Teil nötig, wobei es auf den dazwischenliegenden Teil nicht ankommt. Dementsprechend sind solche Ausführungsformen bevorzugt, die lediglich zwei flankierende Regionen mit mindestens diesen Größen umfassen.

Nach einer alternativen Ausführungsform dieser Verwendung werden Nukleinsäuren mit insgesamt zwei Nukleinsäureabschnitten eingesetzt, die jeweils mindestens 70 bis 150 Nukleinsäurepositionen umfassen und damit den für das Protein codierenden Bereich zumindest teilweise, vorzugsweise vollständig flankieren.

Denn allein um den Austausch der beiden Genkopien über homologe Rekombination zu ermöglichen, braucht es sich dabei nicht zwangsläufig um proteincodierende Abschnitte zu handeln. Vielmehr eignen sich hierfür auch die Randbereiche der betreffenden Gene, welche natürlicherweise eine andere Funktion (Promotor, Terminator, Enhancer etc.) ausüben oder lediglich nichtfunktionelle intergenische Abschnitte darstellen. So kann die funktionelle Inaktivierung beispielsweise auch in der Deletion des Promoters bestehen, wofür es bei einer Deletionsmutation dieser Ausführungsform notwendig ist, auf flankierende, nichtcodierende Abschnitte zurückzugreifen. Je nach Einzelfall kann es auch sinnvoll sein, für die flankierenden Regionen solche Abschnitte auszuwählen, die zum Teil in den proteincodierenden Bereich hineinreichen und zum Teil außerhalb liegen.

Derartige, wenigstens zum Teil nichtcodierende Bereiche können für *B. licheniformis* beispielsweise SEQ ID NO. 31 entnommen werden. Die für die Gene *recA* aus *B. amyloliquefaciens* und für *recA* und *recE* aus *B. subtilis* sind beispielsweise den oben angegebenen Datenbankeinträgen zu entnehmen. Für andere Stämme, beispielsweise auch für *recA* aus *B. licheniformis* ist es möglich, die betreffenden nichtcodierenden Bereiche über PCR-basierte Verfahren aus einer Präparation der genomischen DNA zu erschließen; wie dies in Beispiel 1 veranschaulicht ist. Diese Verfahren (beispielsweise anchored PCR mit nach außen, in einen unbekannten Bereich weisenden Primern) sind im Stand der Technik etabliert. Als Ausgangspunkte hierfür dienen die bekannten Genabschnitte, die dazu dienen, um die noch unbekannten Regionen zu erschließen. Sobald diese nach Amplifizierung sequenziert worden sind, können sie ihrerseits zur Synthese weiterer Primer dienen, und so fort. Der vorliegenden Erfindung zufolge können die hierfür benötigten Primer anhand der SEQ ID NO. 1 und 31 auch für andere Spezies grampositiver Bakterien und hierunter insbesondere für solche der Gattung

Bacillus entworfen werden, gegebenenfalls unter unter Einführung variabler Positionen, wie dies oben bereits erläutert worden ist.

Bei einer entsprechenden Verwendung handelt es sich demzufolge um eine der zuvor beschriebenen erfindungsgemäßen Nukleinsäuren und/oder um eine Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 31 angegebenen Nukleotidsequenz in den Positionen 369 bis 1415 zu mindestens 1045, vorzugsweise mindestens 1046, ganz besonders bevorzugt 1047 dieser 1047 Positionen übereinstimmt, beziehungsweise um die zumindest teilweise nichtcodierenden flankierenden Bereiche zu diesen Nukleinsäuren.

So zeigt beispielsweise ein Vergleich der beiden Nukleinsäuresequenzen von SEQ ID NO. 1 und 31, daß diese sich innerhalb des für das Protein codierenden Bereichs (Positionen 369 bis 1415 gemäß SEQ ID NO. 31) in drei Positionen unterscheiden. Dies sind die Positionen 282, 283 und 284 gemäß SEQ ID NO. 1 (CAC) beziehungsweise 650, 651 und 652 gemäß SEQ ID NO. 31 (ACA). Beide Sequenzen fallen unter den oben als erfindungsgemäß bezeichneten Homologiebereich und kennzeichnen bevorzugte Ausführungsformen des hier dargestellten Erfindungsaspekts: SEQ ID NO. 1 stützt sich dabei auf den kommerziell erhältlichen Stamm DSM 13; SEQ ID NO. 31 wurde durch Nacharbeiten der Erfindung anhand eines prinzipiell beliebigen *B. licheniformis*-Stamms erhalten (Beispiel 1). Da sie in 1044 Positionen übereinstimmen, läßt sich die in den Beispielen als erfolgreich beschriebene Ausführungsform über eine zunehmende Bevorzugung von 1045, 1046 und ganz besonders 1047 übereinstimmenden Positionen zu SEQ ID NO. 31 beschreiben.

Die verschiedenen hierunter fallenden Ausführungsformen sind entsprechend den bisherigen Darstellungen bevorzugt.

In bevorzugten Ausführungsformen dieser Verwendungen handelt es sich bei dem grampositiven Bakterium vorzugsweise um eines der Gattungen *Clostridium* oder *Bacillus* und um eines, das natürlicherweise zur Sporulation befähigt ist, bei dem gleichzeitig mit *recA* ein Gen aus der Phase IV der Sporulation funktionell inaktiviert wird.

Denn viele grampositiven Bakterien sind wie einleitend beschrieben in der Lage, unter entsprechend ungünstigen Umweltbedingungen den Vorgang der Sporulation

einzuleiten. Dies kann erfindungsgemäß insofern für Sicherheitsaspekte ausgenutzt werden, als in Kombination mit der oben dargestellten funktionellen Inaktivierung von *recA* ein Sporulationsgen aus der vergleichsweise späten Phase IV der Sporulation ebenfalls funktionell inaktiviert wird. Damit stehen der formulierten Aufgabe entsprechend zwei gleichzeitig und erfindungsgemäß miteinander kombinierbare Systeme für die Herstellung sicherer GVO zur Verfügung. Die Kombination beider Systeme war bislang noch nicht bekannt, insbesondere nicht zu diesem Zweck.

Besonders erfolgreich konnte die Inaktivierung von Sporulationsgenen an Spezies der Gattungen *Clostridium* und *Bacillus* realisiert werden, weshalb die hierdurch gekennzeichneten Ausführungsformen entsprechend bevorzugt sind. So folgen die in den Beispielen der vorliegenden Anmeldung dargestellten Versuche für *spoV* von *B. licheniformis* konzeptionell demselben molekularbiologischen Vorgehen, wie oben für *recA* beschrieben ist. So konnten gemäß Beispiel 1 die in SEQ ID NO. 19 bis 24 gezeigten Primer erfolgreich zur Gewinnung eines *spoV*-Gens aus einem *B. licheniformis*-Stamm genutzt werden (vergleiche Figur 3A). Beispiel 2 zeigt das Vorgehen zur funktionellen Inaktivierung und Beispiel 3 dessen Erfolg (Figur 4). Die erhofften phänotypischen Sporulationsdefekte sind durch Beispiel 4 und Figur 5 belegt.

Insbesondere ist es überraschend, daß die Verhinderung der Sporulation erst in einem so späten Stadium für diesen Zweck erfolgreich ist. Zwar werden in *spoV*-Mutanten von *B. licheniformis* unter entsprechenden Bedingungen noch Sporen (die sogenannten „Phase-Grau-Sporen“) gebildet, doch sind diese steril und nicht mehr in der Lage auszukeimen. Insofern trägt diese Mutation dem Sicherheitsaspekt Rechnung. Bisher war eine Unterbindung der Sporulation eher zu einem früheren Zeitpunkt favorisiert worden. Die Inaktivierung in Phase IV sorgt zusätzlich jedoch dafür, daß die in den früheren Sporulations-Phasen aktiven Faktoren auch weiterhin in den Mutanten gebildet werden. Ohne an diese Theorie gebunden sein zu wollen, kann man vermuten, daß zumindest einige dieser Faktoren von den Zellen auch für die normalen während der Fermentation ablaufenden Stoffwechselvorgänge benötigt werden. Bei dem Ausschalten zu einem früheren Zeitpunkt stünden sie nicht mehr zur Verfügung. Umgekehrt besteht die vorteilhafte Wirkung der Inaktivierung der Sporulation in Phase IV darin, daß der hierdurch stattfindende Eingriff in die Physiologie der Zellen nicht so gravierend ist und die Fermentation an sich weniger beeinträchtigt wird als bei einem früheren Ausschalten

dieser Gene. Den Erfolg dieses Konzepts zeigt die gemäß Beispiel 5 ermittelte und in Figur 6 dargestellte Wachstumskurve.

Bevorzugt ist eine derartige Verwendung, bei der es sich bei dem inaktivierten Gen aus der Phase IV der Sporulation in der Nomenklatur von *B. subtilis* um eines der Gene *spoIVVA*, *spoIVVB*, *spoIVCA*, *spoIVCB*, *spoIVFA*, *spoIVFB* oder *yqfD* beziehungsweise um ein hierzu homologes Gen handelt, vorzugsweise im Fall von *B. subtilis* um das Gen *yqfD*, im Fall von *Bacillus licheniformis* um das Gen *spoIV* und in jedem anderen Fall um ein hierzu homologes Gen.

All diese Gene sind an sich bekannt und für diese Phase der Sporulation beschrieben. Das *B. subtilis*-Gen *spoIVVA* codiert für das Phase-IV-Sporulationsprotein A, das in den Datenbanken Swiss-Prot (Geneva Bioinformatics (GeneBio) S.A., Genf, Schweiz; <http://www.genebio.com/sprot.html>) und NCBI (siehe oben) unter der Nummer P35149 hinterlegt ist. Es spielt für die Bildung einer intakten Sporenhülle und deren Zusammenbau eine Rolle. Die Aminosäuresequenz des zugehörigen Faktors SpolVA wird in SEQ ID NO. 8 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist des Institute Pasteur, Paris, Frankreich (<http://genolist.pasteur.fr/SubtiList/genome.cgi>) unter der Nummer BG10275 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 7 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 7 angegebene Nukleotidsequenz von 1 bis 1876 erfindungsgemäß als Gen *spoIVVA* bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 1679; das erste Codon, das heißt die Positionen 201 bis 203 werden *in vivo* nicht als Leucin sondern als Methionin translatiert.

Das *B. subtilis*-Gen *spoIVVB* codiert für das Phase-IV-Sporulationsprotein B, das in den Datenbanken Swiss-Prot und NCBI unter der Nummer P17896 hinterlegt ist. Es ist für den Sigma-Faktor-K-abhängigen Übergangspunkt während der Sporulation beziehungsweise dessen Aktivierung in der Mutterzelle von Bedeutung. Es spielt für die interkompartimentelle Signalübertragung eine Rolle, wahrscheinlich über den hydrophoben

N-Terminus. Die Aminosäuresequenz des Faktors SpolVB wird in SEQ ID NO. 10 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10311 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 9 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 9 angegebene Nukleotidsequenz von 1 bis 1675 erfindungsgemäß als Gen *spo/VB* bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 1478.

Das *B. subtilis*-Gen *spo/VCA* codiert für eine putative ortsspezifische DNA-Rekombinase, die in den Datenbanken Swiss-Prot und NCBI unter der Nummer P17867 hinterlegt ist. Sie spielt wahrscheinlich eine Rolle, um die Gene *spoIIIC* und *spo/VCB* zu rekombinieren, woraus der Sigmafaktor K hervorgeht. Die Aminosäuresequenz dieser Rekombinase wird in SEQ ID NO. 12 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10458 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 11 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 11 angegebene Nukleotidsequenz von 1 bis 1900 erfindungsgemäß als Gen *spo/VCA* bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 1703; das erste Codon, das heißt die Positionen 201 bis 203 werden *in vivo* nicht als Valin sondern als Methionin translatiert.

Das *B. subtilis*-Gen *spo/VCB* codiert für den RNA-Polymerase-Sigmafaktor-K-Precursor, der in den Datenbanken Swiss-Prot und NCBI unter der Nummer P12254 hinterlegt ist. Der Rest dieses Faktors wird von dem Gen *spoIIIC* codiert, welches auf dem Chromosom ca. 10 kb entfernt ist, wobei der dazwischenliegende Bereich als SKIN bezeichnet wird. Durch Excision dieses Fragments in der unmittelbar vorangehenden Sporulationsphase wird der aktive Sigma-Faktor K erhalten, welcher seinerseits als

Transkriptionsfaktor wirkt. Die Aminosäuresequenz des Teilstoßers SpoIVCB wird in SEQ ID NO. 14 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10459 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 13 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 13 angegebene Nukleotidsequenz von 1 bis 868 erfindungsgemäß als Gen *spo/VCB* bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 671.

Das *B. subtilis*-Gen *spo/VFA* codiert für das Phase IV-Sporulationsprotein FA. Dieser Faktor, der vermutlich in der Lage ist, mit SpoIVFB (siehe unten) ein Heterodimer zu bilden, erfüllt wahrscheinlich die Aufgabe, diesen Faktor zu stabilisieren aber dadurch gleichzeitig auch zu inhibieren. Deshalb wird SpoIVFA auch schon zu einem früheren Zeitpunkt, vermutlich in Phase II gebildet. Die Aminosäuresequenz von SpoIVFA ist in den Datenbanken Swiss-Prot und NCBI unter der Nummer P26936 hinterlegt. Sie wird in SEQ ID NO. 16 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10331 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 15 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 15 angegebene Nukleotidsequenz von 1 bis 1192 erfindungsgemäß als Gen *spo/VFA* bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 995.

Das *B. subtilis*-Gen *spo/VFB* codiert für das Phase-IV-Sporulationsprotein FB. Dabei handelt es sich um eine membranassoziierte Metalloprotease, die vermutlich für die Prozessierung von Pro-Sigma K zu Sigma K zuständig ist; sie wird ebenfalls bereits in Phase II der Sporulation gebildet. Die Aminosäuresequenz von SpoIVFB ist in den

Datenbanken Swiss-Prot und NCBI unter der Nummer P26937 hinterlegt. Sie wird in SEQ ID NO. 18 der vorliegenden Anmeldung angegeben, und zwar als die vom Programm PatentIn erzeugte Übersetzung der vorangegangenen DNA-Sequenz. Die zugehörige Nukleotidsequenz kann der Datenbank Subtilist unter der Nummer BG10332 entnommen werden und ist im Sequenzprotokoll in SEQ ID NO. 17 angegeben, und zwar zusammen mit den 200 vor dem 5'-Ende und den 197 hinter dem 3'-Ende gelegenen Nukleotiden. Hierbei wird, ungeachtet der Tatsache, daß diese Randsequenzen durchaus sinnvolle genetische Informationen, insbesondere Regulationselemente oder auch Abschnitte anderer Gene enthalten dürften, die komplette unter SEQ ID NO. 17 angegebene Nukleotidsequenz von 1 bis 1264 erfindungsgemäß als Gen *spo/VFB* bezeichnet. Der codierende Bereich erstreckt sich von den Positionen 201 bis 1067; das erste Codon, das heißt die Positionen 201 bis 203 werden *in vivo* nicht als Leucin sondern als Methionin translatiert.

Auch die beiden bevorzugten Gene gehen an sich aus dem Stand der Technik hervor. Die DNA- und Aminosäuresequenzen von *spo/V* aus *B. licheniformis* sind in der Datenbank NCBI unter der Nummer AJ616332 hinterlegt. Dieser Faktor wird in der Diplomarbeit „Arbeiten zur Herstellung einer sporulationsnegativen Mutante von *Bacillus licheniformis*“ von M. Gröne (2002), im Fachbereich Biologie der Westfälischen Wilhelms-Universität Münster, Deutschland als ein für die Sporulation von *B. licheniformis* essentieller Faktor beschrieben. Die zugehörigen, zum Teil auch regulatorische Bereiche umfassenden Sequenzen sind in der vorliegenden Anmeldung unter SEQ ID NO. 3 und 4 angegeben. Zu diesen Sequenzen ist zu bemerken, daß erfindungsgemäß der Bereich der Nukleotide 1 bis 1792 als Gen *spo/V* bezeichnet wird, wobei der eigentliche SpolV-codierende Abschnitt die Positionen 140 bis 1336 umfaßt; die Randsequenzen mögen wiederum andere genetische Elemente wie Regulationselemente oder Teile anderer Gene enthalten. Hierunter wird das erste Codon GTG der Positionen 140 bis 142 *in vivo* nicht als Valin sondern als Methionin translatiert.

In dieser Arbeit wird auch auf den Faktor beziehungsweise das Gen *yqfD* aus *B. subtilis* hingewiesen, welches mit einer Homologie von 68% Identität auf Aminosäureebene als das nächstähnliche bis dato bekannte Protein angesehen wird. Dieser Faktor ist in der Datenbank Swiss-Prot unter der Nummer P54469 angegeben; sowohl die Aminosäuresequenz als auch die DNA-Sequenz mit beiden ca. 200 bp flankierenden Bereichen gehen aus der Datenbank Subtilist unter der Nummer BG11654 hervor. Der

dortige Eintrag vermerkt, es sei zwar ein unbekanntes Protein, aber aufgrund der bestehenden Sequenzhomologien könne es als Ähnliches zum Phase-IV-Sporulationsprotein angesehen werden. Die zugehörigen Sequenzen können SEQ ID NO. 5 und 6 der vorliegenden Anmeldung entnommen werden. Zu diesen Sequenzen ist ergänzend zu bemerken, daß ungeachtet der ebenfalls angegebenen und möglicherweise andere genetische Elemente enthaltenden Randssequenzen erfindungsgemäß der Bereich der Nukleotide 1 bis 1594 als Gen *yqfD* bezeichnet wird, wobei der eigentliche proteincodierende Abschnitt die Positionen 201 bis 1397 umfaßt. Hierunter wird das erste Codon GTG der Positionen 201 bis 203 wie bei *spoV* aus *B. licheniformis* *in vivo* nicht als Valin sondern als Methionin translatiert.

Es ist zu erwarten, daß alle anderen grampositiven, natürlicherweise zur Sporulation befähigten Mikroorganismen über Homologe zu den genannten sieben Genen und davon abgeleitete Faktoren vergleichbarer Funktionen verfügen. Sie dürften in an sich bekannter Weise unter Hybridisierung mit den in im Sequenzprotokoll angegebenen Nukleinsäuren oder wie oben bereits erwähnt über PCR-basierte Ansätze zur Sequenzierung der zugehörigen chromosomalen Abschnitte dieser Mikroorganismen ohne weiteres identifizierbar sein, insbesondere mithilfe der beiden homologen Sequenzen SEQ ID NO. 3 beziehungsweise 5, womit eine gewisse Varianz über Spezies-Grenzen hinweg ermöglicht wird.

Eines dieser Gene, vorzugsweise *yqfD* / *spoV* beziehungsweise dessen Homologes wird im Produktionsstamm erfindungsgemäß gleichzeitig mit *recA* inaktiviert, um hieraus entsprechende Sicherheitsstämme zu erhalten. Vorteilhafterweise werden hierfür entsprechend den oben gemachten Ausführungen zur Deletionsmutagenese die in SEQ ID NO. 3, 5, 7, 9, 11, 13, 15 oder 17 angegebenen Nukleinsäuren selbst zur Inaktivierung verwendet. Damit ist es nicht einmal nötig, die betreffenden homologen Gene aus den für die Produktion verwendeten Spezies selbst zu identifizieren. Hierbei ist zu erwarten, daß diese Deletionen umso erfolgreicher sind, je enger die betreffenden Spezies mit *B. subtilis* beziehungsweise *B. licheniformis* verwandt sind. Denn hiermit sollte eine zunehmende Homologie der betreffenden Gene verbunden sein. Aus diesem Grund sind im Sequenzprotokoll jeweils auch die ca. 200 bp umfassenden Randsequenzen angegeben, denn damit können entsprechend den für *recA* gemachten Ausführungen Konstrukte gebildet werden, die die für ein Crossing-over nötigen mindestens 70 bis 150 Positionen umfassenden Bereiche in vollständig flankierenden

Bereichen enthalten und mit einer gewissen Erfolgswahrscheinlichkeit auch für die Deletion der betreffenden Abschnitte in diesbezüglich nicht näher charakterisierten Mikroorganismen eingesetzt werden können.

Erfindungsgemäß ist es möglich, zusammen mit *recA* mehrere der genannten Phase IV-Gene zu inaktivieren und dadurch Sicherheitsstämme zu erhalten, die neben der Unfähigkeit zur RecA-vermittelten DNA-Rekombination nicht in der Lage sind, reife Sporen zu bilden. Erfindungsgemäß reicht es hierfür jedoch aus, neben *recA* lediglich eines dieser Gene zu inaktivieren, weshalb in einer bevorzugten derartigen Verwendung genau ein Gen aus der Phase IV der Sporulation funktionell inaktiviert wird.

Hierdurch werden Sicherheitsstämme für die biotechnologische Produktion erhalten. Diese sind außerhalb der optimalen Fermentationsbedingungen weniger gut überlebensfähig, insbesondere unter Umweltbedingungen, die schlechte Nährstoffversorgung und DNA-schädigende Einflüsse, etwa durch UV-Strahlung oder aggressive chemische Verbindungen, umfassen. Die genannte erste Gruppe von Umwelteinflüssen würde bei natürlicherweise zur Sporulation befähigten grampositiven Bakterien den Übergang in die Dauerform der Sporen induzieren; die zweite Gruppe von Einflüssen kann von Mikroorganismen natürlicherweise über RecA-vermittelte DNA-Reparatur- und Rekombinationsprozesse ausgeglichen werden. Wenn die Zellen zu beidem nicht mehr oder nur noch stark eingeschränkt in der Lage sind, sind sie erfindungsgemäß als Sicherheitsstämme geeignet.

Ferner ist es möglich, zusammen mit *recA* eines oder mehrere der im Stand der Technik bekannten Gene zu inaktivieren und dadurch Sicherheitsstämme zu erhalten, die neben der Unfähigkeit zur RecA-vermittelten DNA-Rekombination und zur Bildung reifer Sporen auch über diese zusätzlichen Eigenschaften gekennzeichnet sind. Dazu gehören neben „aktiven“, die Lebensfähigkeit unterbindenden und entsprechend stringent zu regulierenden Systemen auch jene, die im einleitend dargestellten Stand der Technik als „passive“ Systeme zur Erzeugung von GVO bezeichnet worden sind. Dazu gehören insbesondere inaktivierende Mutationen in einem oder mehreren der folgenden Gene: *epr*, *rp-I*, *rp-II*, *isp-1*, *apr*, *npr*, *spoOA*, *bpr*, *rsp*, *mpR*, *vpr*, *spoOA*, *spolID*, *spolIAC*, *spo2*, *spo3*, *sigE*, *sigF*, *spolIE*, *spolISB*, *sigG*, *spolVCB*, *spolIIC*, *nprM* und das Gen für die Isopropylmalat-Dehydrogenase (*leuB*). Diese Genbezeichnungen sind dem in der Einleitung zur vorliegenden Anmeldung dargestellten Stand der Technik entnommen.

Somit sind an dieser Stelle mit diesen Abkürzungen jeweils die in den einleitend genannten Anmeldungen und Publikationen beschriebenen Bedeutungen gemeint. Sollten für dieselben Gene beziehungsweise Gengruppen, codierend für dieselben oder homologe Proteine, im Stand der Technik weitere Namen etabliert sein, insbesondere für die Homologen in anderen Bakterienspezies als denen, die den erwähnten Arbeiten zugrundegelegen haben, so gilt das hier Gesagte entsprechend.

Erfindungsgemäß ist es jedoch nicht unbedingt notwendig, neben *recA* und gegebenenfalls zusätzlich einem Sporulationsphasen IV-Gen ein weiteres Gen zu inaktivieren, so daß vorzugsweise auf diese weiteren Mutationen weitgehend verzichtet wird. Hiermit ist der in der Aufgabe zur vorliegenden Anmeldung geforderte Vorteil verbunden, möglichst wenige Sicherheitssysteme parallel in derselben Zelle zu etablieren. Hiermit wird der Arbeitsaufwand geringer gehalten, als wenn man, wie in den genannten Arbeiten zu *B. megaterium* vorgeschlagen, vier verschiedene Deletionen vornehmen müßte. Das ist insbesondere dann relevant, wenn die betreffenden Zellen zuerst - solange sie noch zur Rekombination in der Lage sind - mit den für die Produktion relevanten Transgenen versehen werden und dann erst in Sicherheitsstämme, insbesondere in einen *recA*-Minus-Phänotyp überführt werden. Nur für sehr kritische Fälle, etwa hochpathogene Stämme, sind derartige weitere Mutationen angezeigt.

Aufgrund der mit der vorliegenden Anmeldung zur Verfügung gestellten Sequenzen werden Sporulationsdefekte in den Genen *spo/VA*, *spo/VB*, *spo/VCA*, *spo/VCB*, *spo/VFA*, *spo/VFB* oder *yqfD* in der Nomenklatur von *B. subtilis* beziehungsweise im Fall von *Bacillus licheniformis* in dem Gen *spo/V* beziehungsweise den je nach Wirtszelle vorhandenen hierzu homologen Genen erzeugt.

Diese Homologie läßt sich in erster Näherung über einen Sequenzvergleich erschließen. Zur Kontrolle kann im vorliegenden, für die biotechnologische Produktion vorgesehenen Mikroorganismus-Stamm das fragliche Gen inaktiviert und über eine Wiederherstellung des Phänotyps (Rescue) die funktionelle Übereinstimmung der betreffenden Gene überprüft werden. Überführt die parallele Bereitstellung einer erfindungsrelevanten *spo/VA*, *spo/VB*, *spo/VCA*, *spo/VCB*, *spo/VFA*, *spo/VFB*, *yqfD*- oder *spo/V*-Kopie die betreffende Knock-out-Mutante wieder in einen Sporulations-positiven Phänotyp, so wäre damit der Nachweis erbracht, daß auch eine funktionelle Austauschbarkeit der betrachteten Gene besteht. Unter homologen Genen zu den genannten Phase-IV-

Sporulationsgenen werden erfindungsgemäß also insbesondere solche verstanden, die einem derartigen „Rescue“ zugänglich sind. Wenn das möglich ist, handelt es sich um ein bevorzugt verwendetes Sporulationsgen. Diese Kontrolle ist insbesondere deshalb mit zumutbarem Arbeitsaufwand möglich, weil zum einen erfindungsgemäß eben eine solche funktionell inaktive Mutante erzeugt werden soll und zum anderen über das Sequenzprotokoll zur vorliegenden Anmeldung die betreffenden Sequenzen aus *B. subtilis* und die besonders bevorzugte davon zusätzlich aus *B. licheniformis* zur Verfügung gestellt werden, über die ein derartiger Rescue vorgenommen werden kann.

In bevorzugten Ausführungsformen erfolgt die erfindungsgemäße, bisher beschriebene Verwendung zur funktionellen Inaktivierung der Gene *spo/VA*, *spo/VB*, *spo/VCA*, *spo/VCB*, *spo/VFA*, *spo/VFB*, *yqfD* oder *spo/V* beziehungsweise den hierzu jeweils homologen Genen mithilfe der Sequenzen SEQ ID NO. 3, 5, 7, 9, 11, 13, 15, oder 17 oder Teilen davon, vorzugsweise mithilfe von Teilen, die mindestens 70 bis 150 zusammenhängende Nukleinsäurepositionen umfassen, besonders bevorzugt mithilfe von zwei solchen Teilen, die einen dazwischenliegenden Teil des Gens umschließen.

Denn wie oben beschrieben können über diese konkreten Sequenzen, insbesondere für *B. licheniformis* und *B. subtilis* und nahe mit diesen verwandte Spezies entsprechende molekularbiologische Konstrukte hergestellt werden. Hierfür stehen alle oben zu *recA* ausgeführten Möglichkeiten zur Verfügung und sind entsprechend bevorzugt.

Einen eigenen Gegenstand der vorliegenden Erfindung stellen die mit den beschriebenen Verfahren erhaltenen Mikroorganismen dar. In seiner allgemeinsten Formulierung handelt es sich dabei also um ein grampositives Bakterium, welches nicht *Bacillus megaterium* ist, bei dem das Gen *recA* funktionell inaktiviert ist.

Hierbei ist in erster Linie grampositive Bakterien gedacht, bei denen das Gen *recA* durch gentechnische, das heißt künstliche Arbeitsschritte funktionell inaktiviert worden ist.

Wie bereits gesagt sind grampositive Bakterien, beispielsweise aufgrund ihrer Fähigkeit zur Sekretion von Wertstoffen und/oder ihrer vergleichsweise leichten Fermentierbarkeit die für die Biotechnologie wichtigsten Mikroorganismen. Hierunter werden für die verschiedenen Einsatzgebiete verschiedene Spezies bevorzugt, so werden niedermolekulare Verbindungen wie etwa Aminosäuren in besonders großem Ausmaß mithilfe

von Corynebakterien produziert; *Bacillus* und hierunter insbesondere *B. licheniformis* wird für die Produktion von extrazellulären Proteinen besonders geschätzt. Sie alle sind erfindungsgemäß, zumindest grundsätzlich einer funktionellen Inaktivierung von RecA zugänglich.

Diese erfindungsgemäßen Bakterien zeichnen sich durch die beschriebenen Rekombinationsdefekte aus und weisen deshalb unter natürlichen Bedingungen, insbesondere in Konkurrenz mit anderen Mikroorganismen Nachteile hinsichtlich ihrer Überlebensfähigkeit auf und eignen sich somit als Sicherheitsstämme für die biotechnologische Produktion. Hierbei handelt es sich aus den oben ausgeführten Gründen nicht um *Bacillus megaterium*.

Den bisherigen Ausführungen entsprechend sind darunter solche grampositiven Bakterien bevorzugt, bei denen die funktionelle Inaktivierung über Punktmutagenese, teilweise Deletion oder Insertion oder vollständige Deletion des für das vollständige Protein codierenden Bereichs erfolgt ist.

Den bisherigen Ausführungen entsprechend sind weiterhin darunter solche grampositiven Bakterien bevorzugt, bei denen die funktionelle Inaktivierung über eine erfindungsgemäße für RecA codierende Nukleinsäure und/oder über eine Nukleinsäure erfolgt ist, deren Nukleotidsequenz zu der in SEQ ID NO. 31 angegebenen Nukleotidsequenz in den Positionen 369 bis 1415 zu mindestens 1045, vorzugsweise mindestens 1046, ganz besonders bevorzugt 1047 dieser 1047 Positionen übereinstimmt, beziehungsweise über die zumindest teilweise nichtcodierenden flankierenden Bereiche zu diesen Nukleinsäuren.

Hierbei sind die Nukleinsäuren mit den oben beschriebenen Homologiewerten um SEQ ID NO. 1 oder, wie bereits erläutert, um SEQ ID NO. 31 entsprechend bevorzugt. Mikroorganismen, bei denen die funktionelle Inaktivierung mithilfe der in SEQ ID NO. 1 oder 31 angegebenen Nukleinsäure beziehungsweise Abschnitten davon erfolgt ist, stellen in dieser Hinsicht die am meisten bevorzugten Mikroorganismen dar.

Entsprechend dem oben gesagten sind im Falle einer Mutagenese über Crossing over vorzugsweise Randsequenzen von jeweils mindestens 70 bis 150 bp verwendet worden,

was über eine Sequenzierung der betreffenden chromosomalen Abschnitte überprüft werden kann.

Den bisherigen Ausführungen entsprechend sind weiterhin darunter solche grampositiven Bakterien und hierunter vorzugsweise solche der Gattungen *Clostridium* oder *Bacillus*, bevorzugt, die natürlicherweise zur Sporulation befähigt sind und bei denen gleichzeitig mit *recA* ein Gen aus der Phase IV der Sporulation funktionell inaktiviert ist.

Entsprechend dem oben Gesagten sind hierunter insbesondere solche Gendefekte zu verstehen, die über biotechnologische Arbeitsschritte vorgenommen worden sind.

Den bisherigen Ausführungen entsprechend sind weiterhin darunter solche grampositiven Bakterien bevorzugt, bei denen es sich bei dem inaktivierten Gen aus der Phase IV der Sporulation in der Nomenklatur von *B. subtilis* um eines der Gene *spoIVVA*, *spoIVVB*, *spoIVCA*, *spoIVCB*, *spoIVFA*, *spoIVFB* oder *yqfD* beziehungsweise um ein hierzu homologes Gen handelt, vorzugsweise im Fall von *B. subtilis* um das Gen *yqfD*, im Fall von *Bacillus licheniformis* um das Gen *spoIV* und in jedem anderen Fall um ein hierzu homologes Gen.

In besonderen Fällen, etwa beim Einsatz hochpathogener Stämme können auch mehrere der genannten Sporulationsgene oder eines oder mehrere der im Stand der Technik beschriebenen Gene oder Gengruppen *spoIV/yqfD/Homolog*, *epr*, *rp-I*, *rp-II*, *isp-1*, *apr*, *npr*, *spoOA*, *bpr*, *rsp*, *mpr*, *vpr*, *spoOA*, *spoIID*, *spoIAC*, *spo2*, *spo3*, *sigE*, *sigF*, *spoIIIE*, *spoIIISB*, *sigG*, *spoIVCB*, *spoIIC*, *nprM* und/oder das Gen für die Isopropylmalat-Dehydrogenase (*leuB*) funktionell inaktiviert werden. Unter diesen Abkürzungen sind jeweils die in den einleitend genannten Anmeldungen und Publikationen beschriebenen Bedeutungen zu verstehen, wobei eventuelle Synonyme entsprechend eingeschlossen werden.

Den bisherigen Ausführungen entsprechend handelt es sich dabei jedoch bevorzugt jeweils um ein solches grampositives Bakterium, bei dem – neben der RecA-Inaktivierung – genau ein Gen aus der Phase IV der Sporulation funktionell inaktiviert ist.

Entsprechend dem oben Gesagten ist weiterhin jeweils ein solches derartiges grampositives Bakterium bevorzugt, bei dem die funktionelle Inaktivierung der Gene *spoIVVA*, *spoIVVB*, *spoIVCA*, *spoIVCB*, *spoIVFA*, *spoIVFB*, *yqfD* oder *spoIV* beziehungsweise den hierzu jeweils homologen Genen mithilfe einer der Sequenzen SEQ ID NO. 3, 5, 7, 9, 11, 13, 15, oder 17 oder Teilen davon erfolgt ist, vorzugsweise mithilfe von Teilen, die mindestens 70 bis 150 zusammenhängende Nukleinsäurepositionen umfassen, besonders bevorzugt mithilfe von zwei solchen Teilen, die einen dazwischenliegenden Teil des Gens umschließen.

Dies lässt sich über Präparationen der betreffenden DNA, beispielsweise der chromosomalen DNA eines erfindungsgemäßen Stamms und Restriktionsanalyse oder PCR überprüfen. Als Primer hierfür können die jeweiligen flankierenden Sequenzen verwendet werden, wobei die Größe des PCR-Produkts Aufschluß über das Vorhandensein und gegebenenfalls die Größe von Inserts gibt. Dieses Vorgehen ist beispielhaft für *spoIV* in den Beispielen zur vorliegenden Anmeldung dargestellt.

Entsprechend den bisherigen Ausführungen sind unter den erfindungsgemäßen grampositiven Bakterien besonders solche bevorzugt, bei denen es sich um Vertreter der Gattungen *Clostridium* oder *Bacillus* handelt, insbesondere um solche der Spezies *Bacillus subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, *B. stearothermophilus*, *B. globigii*, *B. clausii* oder *B. lentus*, und ganz besonders um Stämme von *B. licheniformis*.

Einen eigenen Erfindungsgegenstand stellen die Verfahren zur Fermentation eines erfindungsgemäßen grampositiven Bakteriums dar.

Denn diese Verfahren zeichnen sich dadurch aus, daß RecA, vorzugsweise im Zusammenspiel mit den dargestellten bevorzugten Ausführungsformen nicht aktiv ist und der betreffende Stamm im Falle einer versehentlichen Freisetzung in die Umgebung der Anlage ein deutlich minimiertes Sicherheitsrisiko darstellt. Für Verfahren zur Fermentation werden entsprechende Sicherheitsanforderungen gestellt, so daß sie gegebenenfalls nur dann durchführbar sind, wenn sie diese Anforderungen erfüllen.

Daß solch ein Stamm unter den optimalen Bedingungen während der Fermentation nicht grundlegend benachteiligt ist, belegt Beispiel 5 (Figur 6) der vorliegenden Anmeldung; das gilt auch für die dort beschriebene Doppelmutante. Die Inaktivierung von *recA* führt

jedoch zu einer deutlich verringerten Lebensfähigkeit unter UV-Einwirkung. Dabei handelt es sich um einen üblichen Umweltfaktor, mit dem Bakterien bei einem eventuellen Austritt aus der Produktionsanlage in die Umgebung konfrontiert sind. Zudem stellt die UV-Bestrahlung eine übliche Sterilisierungsmethode für Labors und biotechnologische Produktionsstätten dar. Wie die Beispiele ferner belegen, führt die Inaktivierung von *spoV* zu einer drastisch verringerten Sporulationsrate. Beide Ansatzpunkte sind diesen Beispielen zufolge miteinander im selben Bakterienstamm kombinierbar. Zudem ergänzen sich beide „passiven“ Systeme zur Erzeugung von technisch nutzbaren Sicherheitsstämmen aufgrund ihrer unterschiedlichen Wirkprinzipien.

Vorzugsweise handelt es sich dabei um Verfahren zur Herstellung eines Wertstoffs, insbesondere einer niedermolekularen Verbindung oder eines Proteins.

Zwar ist es vorteilhaft, wenn entsprechende Stämme auch im Labormaßstab eingesetzt werden. Doch sind hier die übrigen Rahmenbedingungen im allgemeinen leichter zu erfüllen. Außerdem besteht das Haupeinsatzgebiet der Fermentation von Mikroorganismen in der biotechnologischen Herstellung von Wertstoffen.

Der Bedeutung dieser Wertstoffe entsprechend sind hierunter solche Verfahren bevorzugt, bei denen es sich bei der niedermolekularen Verbindung um einen Naturstoff, einen Nahrungsmittelergänzungsstoff oder um eine pharmazeutisch relevante Verbindung handelt.

Hierunter sind beispielsweise Aminosäuren oder Vitamine zu nennen, die besonders als Nahrungsmittelergänzungsstoffe Verwendung finden. Bei pharmazeutisch relevanten Verbindungen kann es sich um Vor- oder Zwischenstufen zu Medikamenten oder sogar um diese selbst handeln. In all diesen Fällen spricht man auch von Biotransformation, wonach die Stoffwechseleigenschaften der Mikroorganismen ausgenutzt werden, um die ansonsten aufwendige chemische Synthese ganz oder zumindest in einzelnen Schritten zu ersetzen.

In nicht minder bevorzugten Verfahren handelt es sich bei dem Protein um ein Enzym, insbesondere eines aus der Gruppe der α -Amylasen, Proteasen, Cellulasen, Lipasen, Oxidoreduktasen, Peroxidasen, Laccasen, Oxidasen und Hemicellulasen.

Industrielle Enzyme, die mit derartigen Verfahren hergestellt werden, finden beispielsweise in der Nahrungsmittelindustrie Verwendung. So dienen α -Amylasen beispielsweise dazu, um das Altbackenwerden von Brot zu verhindern oder um Fruchtsäfte zu klären. Proteasen werden zum Aufschluß von Proteinen verwendet. All diese Enzyme sind für den Einsatz in Wasch- und Reinigungsmitteln beschrieben, wobei insbesondere die von grampositiven Bakterien bereits natürlicherweise hergestellten Subtilisin-Proteasen einen prominenten Platz einnehmen. Insbesondere in der Textil- und Lederindustrie dienen sie der Aufarbeitung der natürlichen Rohstoffe. Ferner können all diese Enzyme wiederum im Sinne der Biotransformation als Katalysatoren für chemische Reaktionen eingesetzt werden.

Wie einleitend in der Aufgabe formuliert war es erwünscht, ein solches Sicherheitssystem zu finden, das nicht so speziell ist, daß es nicht auch in anderen molekularbiologischen Ansätzen Verwendung finden könnte. Derartige Ansätze können zu einem weiteren eigenständigen Erfindungsgegenstand zusammengefaßt werden.

Das bedeutet allgemein formuliert die Verwendung des oben beschriebenen Faktors RecA und/oder eines RecA, welches mit der in SEQ ID NO. 32 angegebenen Aminosäuresequenz in mindestens 347, vorzugsweise 348 der dort gezeigten 348 Aminosäurepositionen übereinstimmt, in einem molekularbiologischen Reaktionsansatz.

Hierfür werden dessen natürlicherweise vorhandenen, einleitend angegebenen Aktivitäten ausgenutzt.

Dementsprechend bevorzugt ist die Verwendung zum Stabilisieren einzelsträngiger DNA, insbesondere bei einer DNA-Polymerisation, bei *in vitro* erfolgenden Rekombinationsvorgängen, oder zum Überführen doppelsträngiger DNA in einzelsträngige DNA oder umgekehrt.

Denn RecA ist ein DNA-einzelstrangbindendes Protein, welches wie erläutert auch eine gewisse Affinität zu doppelsträngiger DNA aufweist. Bei dem natürlichen Prozeß des Crossing over im Zuge der homologen Rekombination kommt diese Funktion zum Tragen. So kann RecA beispielsweise einer PCR oder einer Präparation von Phagen-DNA zugegeben werden, um die Einzelstränge zu stabilisieren. Wenn *in vitro*

Rekombinationsvorgänge nachvollzogen werden, etwa beim Einführen von Mutationen (so auch bei den oben ausgeführten erfindungsgemäßen Mutationen), kann dies von RecA unterstützt werden. Schließlich ist unter dem Überführen doppelsträngiger DNA in einzelsträngige DNA oder umgekehrt eine Gyrase- oder Gyrase-unterstützende Funktion gemeint. Dies kann für die Beeinflussung der DNA-Topologie, etwa bei der Arbeit mit Plasmid-DNA ausgenutzt werden.

Einen eigenen Erfindungsgegenstand stellen Vektoren dar, die eine zuvor beschriebene, erfindungsgemäße Nukleinsäure enthalten. Denn auch in dieser Form wird die vorliegende Erfindung verwirklicht. So kann diese DNA in Form von Klonierungsvektoren molekularbiologisch bearbeitet oder gelagert werden.

Vorzugsweise handelt es sich bei einem solchen Vektor um einen Expressionsvektor. Denn dieser kann dazu ausgenutzt werden, um ein erfindungsgemäßes RecA herzustellen und den genannten Anwendungen des Faktors zuzuführen.

Einen eigenen Erfindungsgegenstand stellen dementsprechend auch Verfahren zur Herstellung eines zuvor beschriebenen, erfindungsgemäßen Faktors RecA dar.

Vorzugsweise handelt es sich dabei um Verfahren, die unter Einsatz einer der oben beschriebenen Nukleinsäuren, insbesondere solchen mit zunehmenden Hologiewerten zur SEQ ID NO. 1 erfolgen, vorzugsweise eines entsprechenden Expressionsvektors und weiter bevorzugt durch Fermentation eines dieser Nukleinsäure beziehungsweise diesen Expressionsvektor enthaltenden Wirts.

So wird die vorliegende Erfindung dadurch verwirklicht, daß eine Zelle ein solches Gen in Form einer chromosomal Kopie erhält und translatiert. Leichter steuerbar erscheint demgegenüber die Bereitstellung dieses Gens in Form eines Plasmids, das gegebenenfalls in mehreren Kopien für die Bildung dieses Faktors sorgt.

Einen eigenen Erfindungsgegenstand stellt die Verwendung der für einen Faktor RecA codierenden, erfindungsgemäßen Nukleinsäure zur Expression dieses Faktors dar. Denn entsprechend dem oben Gesagten wird dadurch die vorliegende Erfindung zumindest in einem Aspekt verwirklicht.

Vorzugsweise dient das dazu, um diesen Faktor selbst herzustellen, insbesondere in einem oben beschriebenen Verfahren. Alternativ hierzu kann die intrazelluläre Expression auch dazu dienen, um molekularbiologische Aktivitäten der betreffenden Zellen zu modulieren, insbesondere bei *in vivo* erfolgenden Rekombinationsvorgängen.

Hierbei ist beispielsweise an die Inaktivierung durch einen Antisense- oder RNA-Interferenz-Ansatz gedacht, nach welchem die für RecA codierende mRNA gezielt ausgeschaltet oder nur in einem Teil translatierbar macht. Hierdurch kann die Expression dieses Faktors ganz gezielt moduliert werden. Das gilt sowohl für biotechnologische Produktionsstämme als auch für Laboransätze zum Studium molekularbiologischer Aspekte.

Ferner wird die vorliegende Erfindung auch in der Verwendung der für einen Faktor RecA codierenden, oben als erfindungsgemäß beschriebenen Nukleinsäure und/oder einer für einen Faktor RecA codierenden Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 31 angegebenen Nukleotidsequenz in den Positionen 369 bis 1415 zu mindestens 1045, vorzugsweise mindestens 1046, besonders bevorzugt 1047 dieser 1047 Positionen übereinstimmt, zur Inaktivierung dieses Faktors oder des Gens *recA* in einem *In vitro*-Ansatz, insbesondere über Wechselwirkung mit einer zugehörigen Nukleinsäure, verwirklicht.

Das kann besonders für *in-vitro*-Transkriptions- oder-Translationsansätze vorteilhaft sein, um Rekombinationsvorgänge zu unterbinden.

Weitere Ausführungsformen der vorliegenden Erfindung ergeben sich aus dem molekularbiologischen Aspekt, über den die Gene *recA* und *spoV* zugänglich sind. Denn wie in Beispiel 1 dargestellt ist, konnten die zugehörigen DNA-Abschnitte über PCR mithilfe der in SEQ ID NO. 19 bis 30 angegebenen Oligonukleotide aus einem prinzipiell beliebigen *B. licheniformis*-Stamm erhalten werden, so daß die vorliegende Erfindung darüber besonders leicht nacharbeitbar ist.

Eine Ausführungsform der vorliegenden Erfindung ist somit eine für eine Teilsequenz von *recA* codierende oder für eine mit *recA* *in vivo* benachbarte Teilsequenz, vorzugsweise weniger als 1.000 bp, besonders bevorzugt weniger als 500 bp entfernt liegende Nukleinsäure gemäß einer der SEQ ID NO. 25 bis 30.

Denn dabei handelt es sich um Teilsequenzen von recA oder um solche, die möglicherweise über nur wenige hundert dazwischenliegende Basenpaare von recA entfernt liegend. Zunehmend bevorzugt sind 1.000, 900, 800, 700, 600, 500, 400, 300, 200, 100 bp bis hin zu einer unmittelbaren Nachbarschaft, das heißt einer Lage zu Beginn oder zum Ende von recA, vorzugsweise in den gerade noch nicht proteincodierenden Bereichen. Sie können entsprechend Beispiel 1 zur Gewinnung eines recA aus einem dahingehend nicht weiter charakterisierten Stamm erhalten werden, wobei die Erfolgswahrscheinlichkeit mit zunehmender Verwandtschaft zu *B. licheniformis* zunimmt.

Entsprechend den bisherigen Ausführungen und illustriert durch Beispiel 1 entspricht auch eine Verwendung mindestens einer, vorzugsweise von mindestens zwei einander entgegenorientierten Nukleinsäuren gemäß SEQ ID NO. 25 bis 30 zur Amplifizierung eines *in vivo* davon umschlossenen DNA-Bereichs der vorliegenden Erfindung.

Die paarweise Verwendung ergibt sich aus dem PCR-Ansatz, welcher grundsätzlich einander entgegengesetzte Primer verlangt. Die Orientierung der betreffenden Primer kann Figur 3B entnommen werden. Hierbei ist aufgrund der vergleichsweise hohen Homologiewerte grundsätzlich davon auszugehen, daß diese Primer auch in noch nicht charakterisierten recA-Genen eine ähnliche Orientierung einnehmen.

Eine hierunter bevorzugte Verwendungsmöglichkeit besteht darin, zunächst die am weitesten außen bindenden Primer herzustellen (etwa recA6 in Kombination mit recA5 gemäß Figur 3B oder, wenn diese nicht funktionieren sollten, recA1 und/oder recA4), um darüber den dazwischenliegenden Bereich zu erhalten. Dann können zur Herstellung des konkreten Deletionskonstrukts weiter innen bindende Oligonukleotide als Primer eingesetzt werden, beispielsweise recA2 (beispielsweise in Kombination mit recA1) und recA3 (beispielsweise in Kombination mit recA4), wobei die Nukleotidsequenzen der inneren Primer anhand der durch die vorangegangene PCR erhaltenen Sequenzen gegebenenfalls korrigiert werden können. Sollte dieses Vorgehen fehlschlagen, können auch, wie das an sich bekannt und oben bereits gesagt worden ist, PCR-Primer mit Sequenzvariationen eingesetzt werden. Den Erfolg dieses Ansatz bestätigt die Sequenzierung des erhaltenen Fragments, welches zu den in SEQ ID NO. 1 und/oder 31

oder in Figur 2 angegebenen Sequenzen deutliche Homologien aufweisen sollte, wenn es sich wie gewünscht um ein *recA*-Gen handelt.

Vorzugsweise handelt es sich dabei um eine Verwendung zur Amplifizierung eines *recA*-Gens, da damit der Aspekt der Inaktivierung dieses Gens gemäß der vorliegenden Erfindung realisierbar wird.

Dementsprechend bevorzugt handelt es sich also um derartige Verwendungen im Rahmen eines oben eingehend beschriebenen Verfahrens zur funktionellen Inaktivierung des Gens *recA* in einem grampositiven Bakterium, welches nicht *Bacillus megaterium* ist, einschließlich der oben beschriebenen Ausführungsformen dieses Erfindungsaspekts.

In analoger Weise bevorzugt handelt es sich um derartige Verwendungen zur Erzeugung eines grampositiven Bakteriums, welches nicht *Bacillus megaterium* ist, bei dem das Gen *recA* funktionell inaktiviert ist, einschließlich der oben beschriebenen Ausführungsformen dieses Erfindungsaspekts.

Weitere Aspekte der vorliegenden Erfindung sind durch Inaktivierung des Gens *spoV* gekennzeichnet. Entsprechend den zuletzt gemachten Ausführungen zu *recA* handelt es sich bei folgenden Aspekten ebenfalls um Verwirklichungen der vorliegenden Erfindung:

- Eine für eine Teilsequenz von *spoV* codierende oder für eine mit *spoV* *in vivo* benachbarte Teilsequenz, vorzugsweise weniger als 1.000 bp, besonders bevorzugt weniger als 500 bp entfernt liegende Nukleinsäure gemäß einer der SEQ ID NO. 19 bis 24;
- Eine Verwendung mindestens einer, vorzugsweise von mindestens zwei einander entgegenorientierten Nukleinsäuren gemäß SEQ ID NO. 25 bis 30 zur Amplifizierung eines *in vivo* davon umschlossenen DNA-Bereichs;
- eine derartige Verwendung zur Amplifizierung eines *spoV*-Gens;
- eine derartige Verwendung im Rahmen eines Verfahrens zur funktionellen Inaktivierung des Gens *recA* in einem grampositiven Bakterium, welches nicht *Bacillus megaterium* ist, wobei gleichzeitig mit *recA* ein Gen aus der Phase IV der Sporulation funktionell inaktiviert wird, einschließlich der oben beschriebenen Ausführungsformen dieses Erfindungsaspekts;
- eine derartige Verwendung zur Erzeugung eines grampositiven Bakteriums, welches nicht *Bacillus megaterium* ist, vorzugsweise eines der Gattungen *Clostridium* oder

Bacillus, das natürlicherweise zur Sporulation befähigt ist und bei dem gleichzeitig mit *recA* ein Gen aus der Phase IV der Sporulation funktionell inaktiviert ist, einschließlich der oben beschriebenen Ausführungsformen dieses Erfindungsaspekts.

Beispiele

Alle molekularbiologischen Arbeitsschritte folgen Standardmethoden, wie sie beispielsweise in dem Handbuch von Fritsch, Sambrook und Maniatis „Molecular cloning: a laboratory manual“, Cold Spring Harbour Laboratory Press, New York, 1989, oder „Mikrobiologische Methoden: Eine Einführung in grundlegende Arbeitstechniken.“ von E. Bast (1999) Spektrum, Akademischer Verlag, Heidelberg, Berlin, oder vergleichbaren einschlägigen Werken angegeben sind. Enzyme und Baukästen (Kits) wurden nach den Angaben der jeweiligen Hersteller eingesetzt.

Beispiel 1

Isolierung der *spoV*- und der *recA*-Region aus einem *B. licheniformis*-Laborstamm

Zur Umsetzung der vorliegenden Erfindung kann grundsätzlich mit den Sequenzprotokoll angegebenen Sequenzen für *recA* und *spoV* aus *B. licheniformis* DSM 13 (SEQ ID NO. 1 beziehungsweise 3) begonnen werden. Hier wurde noch früher begonnen und zunächst ein PCR-basiertes Verfahren zur Isolierung dieser Gene aus einem *Bacillus*-Stamm angewendet, wie es in der Publikation „A general method for cloning *recA* genes of Gram-positive bacteria by polymerase chain reaction“ (1992) von Duwat et al. in *J. Bacteriol.*, Band 174 (Nr. 15), S. 5171 - 5175, beschrieben ist.

Hierfür wurden anhand der aus verschiedenen grampositiven Bakterien und insbesondere aus *B. licheniformis* DSM 13 zuvor bekannten DNA-Sequenzen der Gene *spoV* und *recA* PCR-Primer synthetisiert, von denen die letztlich erfolgreichen in Tabelle 1 und dem Sequenzprotokoll der vorliegenden Anmeldung aufgeführt sind. Ihre Bindeorte an die jeweiligen Genorte sind in Figur 3 dargestellt.

Tabelle 1: Verwendete Oligonukleotide zur Amplifizierung des *spoV*- sowie des *recA*-Locus.

Bezeichnung	SEQ ID NO.	5'-3'-Sequenz
spo1	19	GGCTGATGCTAACACAGGGCAGTGCATC
spo2	20	CATGAACGGCCTTACGACAGCCA
spo3	21	GTCATAAAACGATTTGCCTGAGG

spo4	22	ATGTTCTGTCCCAGGATTGGCTCCTG
spo6	23	GTTTGACTCTGATCGGAATTCTTGGCG
spo7	24	GCACGAAACGAGCGAGAATGGC
recA1	25	GGAATTCGGCATCAGCTTCACTGGAG
recA2	26	GCTATGTCGACTATACCTTGTATGCGG
recA3	27	GACCTCGGAACAGAGCTGAC
recA4	28	TCAAACGTGCAGTCATTAAGAGAATGGATGG
recA5	29	AAGCTTACGGTTAACGTTCTG
recA6	30	ACACAAACGAATTGAAAGTGTAGCG

Hiermit wurden überlappende Teile sowohl des *spoIV*- als auch des *recA*-Locus mit Hilfe der PCR-Technik aus einer Präparation chromosomaler DNA eines *B. licheniformis*-Laborstamms isoliert. Dieser, als *B. licheniformis* A bezeichnete Stamm diente somit als Beispiel für ein beliebiges grampositives Bakterium. Dieser Ansatz ist in analoger Weise auf prinzipiell alle grampositiven Bakterien anwendbar, insbesondere nachdem diese Primer nun bekannt sind.

Nach Sequenzierung der nach Standard-PCR erhaltenen PCR-Fragmente wurden diese jeweils zu einer Gesamtsequenz zusammengesetzt. Diese stimmte im Fall der *spoIV*-Region über die Gesamtlänge von 1.792 bp zu 100% mit der im Sequenzprotokoll unter SEQ ID NO. 3 angegebenen Sequenz von *B. licheniformis* DSM 13 überein. Deshalb bezeichnet die dortige Spezies-Angabe in Feld <213> nicht den speziellen Stamm DSM 13 oder A sondern allgemein die Spezies.

Ferner sei zu dieser Sequenzdarstellung ergänzt, daß der codierende Bereich die dort gezeigten Positionen 140 bis 1336 (einschließlich des Stop-Codons) umfaßt, wobei die ersten drei für das Startcodon GTG codieren, welches *in vivo* als Methionin translatiert wird. Der gesamte gezeigte Abschnitt von 1.792 bp wird hier als Gen *spoIV* bezeichnet, weil er nicht allein den proteincodierenden Teil sondern auch regulatorische Elemente enthält, die diesem Gen zuzuordnen sind. Diese Genbezeichnung erfolgt auch dessen ungeachtet, daß möglicherweise Abschnitte, die primär anderen Genen zuzurechnen sind, hineinreichen mögen. Dies erscheint deshalb gerechtfertigt, weil Genbereiche manchmal überlappend angeordnet sind.

Im Fall der *recA*-Region wurde eine DNA mit einer Länge von 1.557 bp erhalten, die in dem mit SEQ ID NO. 1 der vorliegenden Anmeldung homologisierbaren, unmittelbar proteincodierenden Bereich in allen bis auf drei Positionen übereinstimmt. Sie ist in SEQ ID NO. 31 angegeben. Die abgeleitete Aminosäuresequenz befindet sich in SEQ ID NO. 32. Die Unterschiede auf DNA-Ebene zu SEQ ID NO. 1 liegen in den Positionen 282-284, wodurch die zugehörigen Codons statt für die Aminosäureabfolge DT für EH kodieren. Wegen dieser Abweichung wurden die Spezies-Angaben zu SEQ ID NO. 1 und 31 im jeweiligen Feld <213> um die Stammbezeichnungen DSM 13 beziehungsweise A ergänzt. Es sei hinzugefügt, daß der codierende Bereich die in SEQ ID NO. 31 gezeigten Positionen 369 bis 1415 (einschließlich des Stop-Codons) umfaßt. Entsprechend der Erläuterungen zu *spoV* wird der gesamte gezeigte Abschnitt von 1.557 bp als Gen *recA* bezeichnet.

Die beiden auf diese Weise erhaltenen Loci *spoV* und *recA* wurden bei der Datenbank GenBank (National Center for Biotechnology Information NCBI, National Institutes of Health, Bethesda, MD, USA) unter den Zugangsnummern AJ616332 (für *spoV*) und AJ511368 (für *recA*) hinterlegt.

Beispiel 2

Deletion des *spoV*- und des *recA*-Gens durch gezielte Gendisruption

Mit Hilfe der Technik der gezielten Gendisruption sollte anschließend jeweils ein möglichst großer Bereich aus dem *spoV*- beziehungsweise dem *recA*-Gen deletiert werden. Der Versuchsaufbau ist in Figur 3 skizziert. Teil A zeigt die Einführung der Deletion in *spoV* und damit die Ableitung des Stamms *B. licheniformis* A.1 (Δ *spoV*) aus *B. licheniformis* A. Teil B zeigt die Weiterentwicklung von *B. licheniformis* A.1 (Δ *spoV*) zu *B. licheniformis* A.2 (Δ *spoV*, Δ *recA*). Der betreffende Genort, einschließlich der jeweils unmittelbar flankierenden Gene ist ebenso bezeichnet wie wichtige Restriktionsschnittstellen und die Bindebereiche für die in Beispiel 1 aufgelisteten Primer.

Für die Deletionen wurden, wie unten noch näher erläutert wird, mit den in Figur 3 gezeigten Oligonukleotiden Flankenbereiche aus der chromosomal DNA amplifiziert und zur Konstruktion entsprechender Deletionskartuschen verwendet. Diese wurden zunächst im *E. coli*-Vektor pUCBM21 erstellt. Dieser ist unter

http://seq.yeastgenome.org/vectordb/vector_descrip/PUCBM21.html beschrieben (eingesehen am 14.1.2005) und von der Firma, Roche Diagnostics GmbH, Roche Applied Science, Sandhofer Str. 116, 68305 Mannheim (ehemals Boehringer) kommerziell erhältlich. Später wurden sie in den *Bacillus*-Vektor pE194 umkloniert. Dieser ist unter http://seq.yeastgenome.org/vectordb/vector_descrip/PE194.html beschrieben (eingesehen am 14.1.2005) und von der American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110-2209, USA (<http://www.atcc.org>) erhältlich.

Die Erzeugung einer Gendisruption mit Hilfe solcher integrativer Vektoren erfolgt durch Rekombinationsereignisse über die entsprechenden homologen Flankenbereiche. Dabei wird durch zwei aufeinanderfolgende Einzel-Crossover-Ereignisse die ursprünglich Plasmid-lokalierte, *in vitro* mutierte Kopie des zu disruptierenden Gens gegen die native, intakte Kopie im Bakterienchromosom ausgetauscht. Da der *Bacillus*-Vektor einen temperatursensitiven Replikationsursprung trägt, lassen sich die Plasmidanteile nach der erfolgten Disruption unter nicht-permissiven Bedingungen (42°C) später wieder aus den Zellen entfernen, was die Etablierung einer stabilen Mutantenlinie ermöglicht.

Zur Konstruktion der *spo/V*-Deletionskartusche wurden die Oligonukleotide *spo3* und *spo4* sowie *spo7* und *spo6* verwendet; beide Flanken sind jeweils ca. 450 bp groß und umrahmen einen Bereich von 740 bp (Größe der späteren Deletion). Zur Konstruktion der *recA*-Deletionskartusche wurden die Oligonukleotide *recA1* und *recA2* sowie *recA3* und *recA4* verwendet; beide Flanken sind ca. 340 bp groß und umrahmen einen Bereich von 852 bp (Größe der späteren Deletion). Nach Umlonierung der Deletionskartuschen in die singuläre *PstI*-Schnittstelle von pE194 (durchgeführt in *B. subtilis* DB104; Stamm beschrieben in Kawamura, F. und Doi, R. H. (1984), *J. Bacteriol.*, Band 160, Seiten 442-444) wurden die beiden Disruptionsvektoren pESpo2 sowie pErecA2 erhalten. Zunächst wurde für eine gezielte Deletion des *spo/V*-Gens der Vektor pESpo2 in *B. licheniformis* A via Protoplastentechnik (beschrieben in S. Chang und S.N. Cohen, (1979) *Molec. Gen. Genet.*, Band 168, Seiten 111-115) transformiert. Aus einer entsprechenden Transformantenlinie wurde anschließend unter nicht-permissiven Bedingungen der Vektor wieder aus den Zellen ausgedünnt. Gleichzeitig wurde mittels PCR unter Verwendung der Oligonukleotide *spo1* und *spo2* auf das Vorhandensein eines Mutantenamplifikats durchsucht und nach mehreren Kultivierungspassagen eine stabile Δ *spo/V*-Mutantenlinie (bezeichnet mit *B. licheniformis* A.1) erfolgreich isoliert.

Diese Mutantenlinie A.1 wurde für eine weitere Transformation mit dem *recA*-Disruptionsvektor pErecA2 herangezogen. In analoger Weise wurde auch hier über mehrere Kultivierungspassagen die Transformante bei 42°C subkultiviert, wobei eine entsprechende Δ *recA*-Mutante (Δ *spoV*/ Δ *recA*-Doppelmutante, als *B. licheniformis* A.2 bezeichnet) mit Hilfe eines Screenings auf Mitomycin C-Sensitivität (0,03 µg/µl) identifiziert werden konnte. Zusätzlich wurde dieser phänotypische Befund durch eine PCR unter Verwendung der Oligonukleotide *recA6* und *recA5* verifiziert.

Beispiel 3

Genotypische Charakterisierung der Δ *spoV*-Einzel- und der Δ *spoV*/ Δ *recA*-Doppelmutante

Beide Mutantenstämme (*B. licheniformis* A.1 und A.2) wurden im Vergleich mit dem Ausgangsstamm *B. licheniformis* A auf DNA-Ebene untersucht, um die Deletionen (Verkürzungen) im entsprechenden Genbereich zu überprüfen. So konnte mit Hilfe der PCR-Technik unter Verwendung der Primer *spo1* und *spo2* die 740 bp-Deletion im *spoV*-Locus der Stämme A.1 und A.2 eindeutig nachgewiesen werden (Figur 4 A linker Teil). Gleches gilt für die 852 bp-Deletion im *recA*-Gen der Mutante A.2 unter Verwendung des Primerpaars *recA6* und *recA5* (Figur 4 A, rechter Teil).

Desweiteren wurden die drei Stämme einer *Southern*-Analyse unterzogen. Für den Nachweis der *spoV*-Deletion wurden jeweils 2 µg chromosomaler DNA mit der Restriktionsendonuklease *Cla*I geschnitten und nach gelelektrophoretischer Auftrennung nach Standardmethoden mit einem DIG-markierten PCR-Produkt (erstellt mit den Primern *spo3* und *spo4* anhand der Ausgangs-DNA) hybridisiert. Dabei erschien das größere der beiden detektierten *Cla*I-Fragmente sowohl beim Stamm A.1 als auch beim A.2 auf einer Höhe, die um eine der Deletion entsprechende Größe niedriger als beim Ausgangsstamm A lag (Figur 4 B linker Teil).

Für den Nachweis der *recA*-Deletion wurde die DNA jeweils mit der Restriktionsendonuklease *Ssp*I verdaut und in analoger Weise mit einem DIG-markiertem PCR-Produkt (in analoger Weise mit den Primern *recA1* und *recA2* erstellt) hybridisiert. Hier lag das entsprechende *Ssp*I-Fragment beim Stamm A.2 aufgrund der

Deletion ebenfalls bei einer entsprechend geringeren Höhe als im Parentalstamm A.1 beziehungsweise dem Wildtypstamm A (Figur 4 B rechter Teil).

Beispiel 4

Phänotypische Charakterisierung der $\Delta spoIV$ -Einzelmutante A.1 und der $\Delta spoIV\Delta recA$ -Doppelmutante A.2: Überlebensrate und Sporenbildung

Die Anzucht für die Sporulationstests erfolgte in 200 ml-Schaeffer'schem Sporulationsmedium (16,0 g LB-Medium, 2,0 g KCl, 0,5 g $MgSO_4 \times 7 H_2O$, ad 993,0 ml dest. Wasser; pH 7,0; die Lösung wird autoklaviert und anschließend mit nachfolgenden Komponenten supplementiert: 1 ml $Ca(NO_3)_2$ (0,1 M), 1 ml $MnCl_2$ (0,1 M), 1 ml $FeSO_4$ (1 mM), 4 ml Glucose (20 % (w/v)), in 500 ml Zwei-Schikane-Kolben. Für jeden der drei zu testenden Stämme wurden je drei Kolben 0,25 %ig aus einer LB-Vorkultur angeimpft und bei 30°C sowie ca. 120 rpm (Fa. Innova 4230, New Brunswick Scientific, Edison, NY, USA) inkubiert. Zum Zeitpunkt der Probennahmen wurden jeweils 1.100 μl Kultur in ein steriles Eppendorfgefäß überführt. 100 μl dieser Aliquots wurden zur Bestimmung der Lebendzellzahl eingesetzt, indem eine Verdünnungsreihe in 15 mM NaCl angelegt wurde. Die jeweiligen Verdünnungsstufen wurden auf je vier LB-Agarplatten ausplattiert und bei 30°C über Nacht inkubiert. Die Anzahl lebensfähiger Zellen wurde durch Auszählen der Kolonien auf jeder der vier Agarplatten bestimmt. Die Lebendzellzahl [*colony forming units (cfu)*] wurde unter Berücksichtigung des ausplattierten Volumens und der Verdünnungsstufe ermittelt und die Werte eines Plattensatzes gemittelt. Die restlichen 1.000 μl -Proben wurden für 30 min bei 80°C im Wasserbad inkubiert. Jeweils 250 μl der so behandelten Suspension wurden auf vier LB-Agarplatten ausplattiert und bei 30°C inkubiert. Der Sporentiter wurde durch Auszählen der ausgekeimten Sporen, welche eine Einzelkolonie bilden, ermittelt. Die Sporenanzahl der Platten wurde gemittelt und anschließend pro ml Kultur berechnet. Die auf diese Weise erhaltenen Ergebnisse sind in Tabelle 2 und Figur 5 dargestellt.

Tabelle 2: Mittelwerte der Lebendzellzahl und der überlebenden Sporen pro ml Kultur.

Jeder Stamm wurde in drei parallelen Experimenten (=Anzuchten) untersucht. Jedes Experiment wurde durch Vierfachbestimmung statistisch abgesichert.

A = Ausgangsstamm *B. licheniformis* A; A.1 = *B. licheniformis* A.1 ($\Delta spoV$); A.2 = *B. licheniformis* MD1.2 ($\Delta spoV$, $\Delta recA$).

Zeit [h]	A		A.1		A.2	
	Lebendzellzahl [cfu/ml]	Sporen / ml	Lebendzellzahl [cfu/ml]	Sporen/ ml	Lebendzellzahl [cfu/ml]	Sporen / ml
1	1,9E+04	0	2,6E+04	0	2,3E+04	0
3	-	0	-	0	-	0
6	1,7E+05	0	6,5E+04	0	1,7E+05	0
9	2,3E+06	0	1,2E+05	0	1,2E+06	0
12	3,6E+06	0	2,3E+05	0	2,3E+07	0
24	2,6E+07	0	1,6E+07	0	9,6E+07	0
36	2,7E+08	0,125	3,6E+08	0	3,6E+08	0
48	-	1,375	-	0	-	0
72	1,1E+09	13,5	1,6E+09	0	2,7E+09	0
96	2,6E+09	18,25	1,8E+09	0	3,9E+09	0
168	2,2E+09	33,1	1,5E+09	0	1,3E+09	0
216	6,7E+06	35	6,5E+05	0	9,4E+06	0
264	5,5E+05	36	10	0	3,7E+03	0
336	-	34	-	0	-	0

Man erkennt anhand dieser Ergebnisse, daß lediglich die Zellen des Ausgangsstamms zur Sporenbildung in der Lage sind. Die beiden Mutanten können dies aufgrund der Deletion von *spoV* nicht mehr. Dabei zeigt die Mutante, die zusätzlich durch Deletion des Gens *recA* gekennzeichnet ist, einen etwas weniger steilen Abfall in der Lebendzellzahl.

Beispiel 5

Wachstumskurven der $\Delta spoV$ -Einzelmutante A.1 und der $\Delta spoV/\Delta recA$ -Doppelmutante A.2

Für diesen Test wurden zunächst 10 ml-Kulturen der drei zuvor beschriebenen Stämme in LB-Medium mit jeweils einer Einzelkolonie (von LB-Platte) inkuliert und über Nacht

bei 37°C und 150 rpm inkubiert. Mit diesen Vorkulturen wurden jeweils 50 ml Minimalmedium nach Sambrook et al. 1989 (1% (w/v) Glucose, 0,1 mM CaCl₂, 0,01% (w/v) Hefe-Extrakt, 0,02% (w/v) Casaminoacids; pH 7,0) in 250 ml-Zweischikane-Erlenmeyerkolben 2%ig angeimpft. Diese Kulturen wurden bei 37°C und 180 rpm (Schüttler Innova 4230, New Brunswick Scientific, Edison, NY, USA) bis zum Erreichen einer OD₅₄₆ von ca. 1,0 (späte logaritische Wachstumsphase) angezogen.

Die erhaltene Wachstumskurve bis zum Erreichen der exponentiellen Wachstumsphase ist in Figur 6 dargestellt. Man erkennt, daß alle drei Stämme hinsichtlich ihrer Verdopplungsrate praktisch gleiche Ergebnisse liefern. Ein Wechsel von einem Stamm auf den anderen geht also mit keiner detektierbaren Beeinträchtigung des Wachstums einher.

Beispiel 6

Phänotypische Charakterisierung der Δ spo/V-Einzelmutante A.1 und der Δ spo/V/ Δ recA-Doppelmutante A.2: UV-Sensitivität

Im Anschluß an das vorangegangene Beispiel wurden die Zellen zu dem in Figur 6 gezeigten Zeitpunkt der logarithmischen Wachstumsphase für einen quantitativen Vergleich ihrer UV-Sensitivität jeweils mit 15 mM NaCl-Lösung bis zu einer Stufe von 10⁻⁴ verdünnt und 100 µl Aliquots der Verdünnungen auf LB-Platten ausplattiert. Mit Ausnahme von jeweils zwei LB-Platten, welche später als „Kontrollplatten“ dienten, wurden die restlichen LB-Platten anschließend unterschiedlich lange mit UV-Licht der Wellenlänge 254 nm bestrahlt: die Platten wurden unter einer UV-Lampe, deren Leistung 100 µW/cm² betrug, plaziert, nach verschiedenen Bestrahlungszeiten abgedeckt und zum Schutz vor weiterer Lichteinwirkung in Aluminiumfolie eingewickelt. Dabei entsprachen die unterschiedlichen Bestrahlungszeiten von 2 bis 60 s bei obengenannter Leistung der Lampe UV-Strahlungsintensitäten von 2 bis 60 J/m². „Kontroll-“ und „UV-Platten“ wurden anschließend für 16 h bei 37°C im Dunkeln inkubiert. Anhand der Koloniezahlen auf den Kontrollplatten, deren Werte einer Überlebensrate von 100% entspricht, und auf den UV-Platten wurden die prozentualen Überlebensraten in Abhängigkeit von der UV-Strahlungsintensität berechnet. Die Doppelbestimmungen aus insgesamt drei Tests (=Anzuchten) wurden gemittelt (siehe Tabelle 3) und in Form eines Kurvendiagramms dargestellt (Figur 7A).

Tabelle 3: Gemittelte, prozentuale Überlebensraten der Stämme A, A.1 und A.2 in Abhängigkeit von der UV-Strahlungsintensität.

Jeder Stamm wurde in drei Experimenten (=Anzuchten) untersucht. Jedes Experiment wurde durch Doppelbestimmung statistisch abgesichert. A = *B. licheniformis* A; A.1 = *B. licheniformis* A.1 (Δ spo/V); A.2 = *B. licheniformis* A.2 (Δ spo/V, Δ recA).

UV-Strahlungsintensität (J/m ²)	prozentuale Überlebensrate (%)		
	A	A.1	A.2
0	100	100	100
2	---	---	7,73
4	---	---	3,28
6	---	---	0,87
8	---	---	0,28
10	46,7	43,1	0,01
20	18,73	18,85	0,01
30	6,27	6,02	0,01
40	1,95	2,08	0,01
50	0,33	0,36	0,01
60	0,03	0,13	0,01

Für einen qualitativen Vergleich der beiden Stämme A.1 und A.2 hinsichtlich ihrer UV-Sensitivität wurden jeweils 15 μ l-Aliquots aus Parallelanzuchten auf vier LB-Platten ausgestrichen. Anschließend wurden die Platten zur Hälfte mit einer Plexiglasscheibe abgedeckt (rechte Seite) und auf der anderen Hälfte (linke Seite) mit UV-Licht bestrahlt. Anschließend wurden sie ähnlich wie oben für 16 h bei 37°C im Dunkeln inkubiert.

Das mit diesem Versuch erhaltene Ergebnis ist in Figur 7B gezeigt. Man erkennt, daß die untersuchte Doppelmutante erheblich UV-sensitiver als die Einfachmutante ist. Beide Versuchsteile belegen also, daß die Deletion von *recA*, insbesondere in Kombination mit *spo/V* zu unter Umwelteinflüssen in der freien Natur nicht überlebensfähigen Mutanten führt. Dieser Effekt kann wie in der Beschreibung dargestellt zur Verwendung dieser Mutanten als Sicherheitsstämme genutzt werden.

Beschreibung der Figuren

Figur 1: Aminosäuresequenz-Alignment von SEQ ID NO. 2 mit denen der im Stand der Technik beschriebenen nächstähnlichen Rec-Faktoren.

Dabei bedeuten:

- 1: Faktor RecA aus *B. licheniformis* DSM 13 (SEQ ID NO. 2)
- 2: Faktor RecA aus *B. amyloliquefaciens* (AJ515542 in NCBI)
- 3: Faktor RecA aus *B. subtilis* (Z99112 in NCBI; Region 161035 bis 162078)
- 4: Faktor RecE aus *B. subtilis* (X52132 in NCBI)

Figur 2: Nukleinsäuresequenz-Alignment von SEQ ID NO. 1 mit denen der im Stand der Technik beschriebenen nächstähnlichen *rec*-Gene.

Dabei bedeuten:

- 1: Gen *recA* aus *B. licheniformis* DSM 13 (SEQ ID NO. 1)
- 2: Gen *recA* aus *B. amyloliquefaciens* (AJ515542 in NCBI)
- 3: Gen *recA* aus *B. subtilis* (Z99112 in NCBI; Region 161035 bis 162078)
- 4: Gen *recE* aus *B. subtilis* (X52132 in NCBI)

Figur 3: Schematische Darstellung der genetischen Organisationen der Wildtyp- sowie der Mutanten-Loci von *spoIV* (A) und *recA* (B), einschließlich der Bindestellen für die unter SEQ ID NO. 19 bis 30 angegebenen Primer.

- A Funktionelle Inaktivierung (Deletion) von *spoIV*, das heißt Ableitung des Stamms *B. licheniformis* A.1 aus *B. licheniformis* A (siehe Beispiel 2).
- B Funktionelle Inaktivierung (Deletion) von *recA*, das heißt Ableitung des Stamms *B. licheniformis* A.2 aus *B. licheniformis* A.1 (siehe Beispiel 2).

Figur 4: Genotypische Untersuchung der Mutantenstämme A.1 und A.2 im Vergleich mit dem Ausgangsstamm *B. licheniformis* A mittels PCR (A) und Southern-Analyse (B) (siehe Beispiel 3).

Figur 5: Graphische Auftragung der Lebendzellzahlen sowie des Sporenitzers der *B. licheniformis*-Anzuchten. Jede Kultur wurde in drei parallelen Experimenten untersucht. Jedes Experiment wurde durch Vierfachbestimmung statistisch abgesichert (siehe Beispiel 4).

Dabei bedeuten:

schwarzes, ausgefülltes Quadrat: *B. licheniformis* A;
nicht-ausgefüllter Kreis: *B. licheniformis* A.1 ($\Delta spoIV$);
nicht-ausgefülltes Dreieck: *B. licheniformis* A.2 ($\Delta spoIV$, $\Delta recA$);
gestrichelte Kurven: Lebendzellzahlen
durchgezogene Kurven: jeweiliger Sporeniter

Figur 6: Wachstumskurve einer Anzucht der drei *B. licheniformis*-Stämme in Minimalmedium (siehe Beispiel 5).

Figur 7: Ergebnisse der UV-Tests.

A Graphische Auftragung der Überlebensraten nach UV-Bestrahlung. Jede Kultur wurde in drei parallelen Experimenten untersucht. Jedes Experiment wurde durch Doppelbestimmung statistisch abgesichert (siehe Beispiel 6).

Dabei bedeuten:

schwarzes, ausgefülltes Quadrat: *B. licheniformis* A
nicht-ausgefüllter Kreis: *B. licheniformis* A.1 ($\Delta spoIV$)
nicht-ausgefülltes Dreieck: *B. licheniformis* A.2 ($\Delta spoIV$, $\Delta recA$).

B Qualitativer UV-Test mit Ausstrichen der Stämme A.1 und A.2 auf Platten, welche während der Bestrahlung halb abgedeckt wurden.

Patentansprüche

1. Faktor RecA mit einer Aminosäuresequenz, die zu der in SEQ ID NO. 2 angegebenen Aminosäuresequenz mindestens zu 96% identisch ist.
2. Faktor nach Anspruch 1 mit einer Aminosäuresequenz, die zu der in SEQ ID NO. 2 angegebenen Aminosäuresequenz zunehmend bevorzugt mindestens zu 96,5%, 97%, 97,5%, 98%, 98,5%, 99%, 99,5% und ganz besonders bevorzugt zu 100% identisch ist.
3. Faktor RecA, codiert von einer Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist.
4. Faktor nach Anspruch 3, codiert von einer Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz zunehmend bevorzugt mindestens zu 87,5%, 90%, 92,5%, 95%, 96%, 97%, 98%, 99% und ganz besonders bevorzugt zu 100% identisch ist.
5. Für einen Faktor RecA codierende Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz mindestens zu 85% identisch ist.
6. Nukleinsäure nach Anspruch 6, deren Nukleotidsequenz zu der in SEQ ID NO. 1 angegebenen Nukleotidsequenz zunehmend bevorzugt mindestens zu 87,5%, 90%, 92,5%, 95%, 96%, 97%, 98%, 99% und ganz besonders bevorzugt zu 100% identisch ist.
7. Nukleinsäure nach Anspruch 5 oder 6, codierend für einen Faktor RecA nach einem der Ansprüche 1 bis 4.
8. Verwendung einer für einen Faktor RecA codierenden Nukleinsäure zur funktionellen Inaktivierung des Gens *recA* in einem grampositiven Bakterium, welches nicht *Bacillus megaterium* ist.

9. Verwendung nach Anspruch 8, wobei eine für ein nichtaktives Protein codierende Nukleinsäure mit einer Punktmutation eingesetzt wird.
10. Verwendung nach Anspruch 8, wobei eine Nukleinsäure mit einer Deletions- oder Insertionsmutation eingesetzt wird, vorzugsweise umfassend die jeweils mindestens 70 bis 150 Nukleinsäurepositionen umfassenden Randsequenzen des für das Protein codierenden Bereichs.
11. Verwendung nach Anspruch 8, wobei Nukleinsäuren mit insgesamt zwei Nukleinsäureabschnitten eingesetzt werden, die jeweils mindestens 70 bis 150 Nukleinsäurepositionen umfassen und damit den für das Protein codierenden Bereich zumindest teilweise, vorzugsweise vollständig flankieren.
12. Verwendung nach einem der Ansprüche 8 bis 11, wobei es sich um eine Nukleinsäure nach einem der Ansprüche 5 bis 7 und/oder um eine Nukleinsäure handelt, deren Nukleotidsequenz zu der in SEQ ID NO. 31 angegebenen Nukleotidsequenz in den Positionen 369 bis 1415 zu mindestens 1045, vorzugsweise mindestens 1046, ganz besonders bevorzugt 1047 dieser 1047 Positionen übereinstimmt, beziehungsweise um die zumindest teilweise nichtcodierenden flankierenden Bereiche zu diesen Nukleinsäuren.
13. Verwendung nach einem der Ansprüche 8 bis 12, wobei das grampositive Bakterium, vorzugsweise eines der Gattungen *Clostridium* oder *Bacillus*, natürlicherweise zur Sporulation befähigt ist und gleichzeitig mit *recA* ein Gen aus der Phase IV der Sporulation funktionell inaktiviert wird.
14. Verwendung nach Anspruch 13, wobei es sich bei dem inaktivierten Gen aus der Phase IV der Sporulation in der Nomenklatur von *B. subtilis* um eines der Gene *spoIVVA*, *spoIVVB*, *spoIVCA*, *spoIVCB*, *spoIVFA*, *spoIVFB* oder *yqfD* beziehungsweise um ein hierzu homologes Gen handelt, vorzugsweise im Fall von *B. subtilis* um das Gen *yqfD*, im Fall von *Bacillus licheniformis* um das Gen *spoIV* und in jedem anderen Fall um ein hierzu homologes Gen.

15. Verwendung nach Anspruch 13 oder 14, wobei genau ein Gen aus der Phase IV der Sporulation funktionell inaktiviert wird.
16. Verwendung nach Anspruch 14 oder 15, wobei die funktionelle Inaktivierung der Gene *spoIVVA*, *spoIVVB*, *spoIVCA*, *spoIVCB*, *spoIVFA*, *spoIVFB*, *yqfD* oder *spoIV* beziehungsweise den hierzu jeweils homologen Genen mithilfe der Sequenzen SEQ ID NO. 3, 5, 7, 9, 11, 13, 15, oder 17 oder Teilen davon erfolgt, vorzugsweise mithilfe von Teilen, die mindestens 70 bis 150 zusammenhängende Nukleinsäurepositionen umfassen, besonders bevorzugt mithilfe von zwei solchen Teilen, die einen dazwischenliegenden Teil des Gens umschließen.
17. Grampositives Bakterium, welches nicht *Bacillus megaterium* ist, bei dem das Gen *recA* funktionell inaktiviert ist.
18. Grampositives Bakterium nach Anspruch 17, wobei die funktionelle Inaktivierung über Punktmutagenese, teilweise Deletion oder Insertion oder vollständige Deletion des für das vollständige Protein codierenden Bereichs erfolgt ist.
19. Grampositives Bakterium nach Anspruch 17 oder 18, wobei die funktionelle Inaktivierung über eine Nukleinsäure nach einem der Ansprüche 5 bis 7 und/oder über eine Nukleinsäure erfolgt ist, deren Nukleotidsequenz zu der in SEQ ID NO. 31 angegebenen Nukleotidsequenz in den Positionen 369 bis 1415 zu mindestens 1045, vorzugsweise mindestens 1046, ganz besonders bevorzugt 1047 dieser 1047 Positionen übereinstimmt, beziehungsweise über die zumindest teilweise nichtcodierenden flankierenden Bereiche zu diesen Nukleinsäuren.
20. Grampositives Bakterium nach einem der Ansprüche 17 bis 19, vorzugsweise eines der Gattungen *Clostridium* oder *Bacillus*, das natürlicherweise zur Sporulation befähigt ist und bei dem gleichzeitig mit *recA* ein Gen aus der Phase IV der Sporulation funktionell inaktiviert ist.
21. Grampositives Bakterium nach Anspruch 20, wobei es sich bei dem inaktivierten Gen aus der Phase IV der Sporulation in der Nomenklatur von *B. subtilis* um eines der Gene *spoIVVA*, *spoIVVB*, *spoIVCA*, *spoIVCB*, *spoIVFA*, *spoIVFB* oder *yqfD*

beziehungsweise um ein hierzu homologes Gen handelt, vorzugsweise im Fall von *B. subtilis* um das Gen *yqfD*, im Fall von *Bacillus licheniformis* um das Gen *spoIV* und in jedem anderen Fall um ein hierzu homologes Gen.

22. Grampositives Bakterium nach Anspruch 20 oder 21, wobei genau ein Gen aus der Phase IV der Sporulation funktionell inaktiviert ist.
23. Grampositives Bakterium nach Anspruch 21 oder 22, wobei die funktionelle Inaktivierung der Gene *spoIV*_A, *spoIV*_B, *spoIV*_C_A, *spoIV*_C_B, *spoIV*_F_A, *spoIV*_F_B, *yqfD* oder *spoIV* beziehungsweise den hierzu jeweils homologen Genen mithilfe der Sequenzen SEQ ID NO. 3, 5, 7, 9, 11, 13, 15, oder 17 oder Teilen davon erfolgt ist, vorzugsweise mithilfe von Teilen, die mindestens 70 bis 150 zusammenhängende Nukleinsäurepositionen umfassen, besonders bevorzugt mithilfe von zwei solchen Teilen, die einen dazwischenliegenden Teil des Gens umschließen.
24. Grampositives Bakterium nach einem der Ansprüche 17 bis 23, wobei es sich um eines der Gattungen *Clostridium* oder *Bacillus* handelt, insbesondere um eines der Spezies *Bacillus subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, *B. stearothermophilus*, *B. globigii*, *B. clausii* oder *B. lentus*, und ganz besonders um einen Stamm von *B. licheniformis*.
25. Verfahren zur Fermentation eines grampositiven Bakteriums nach einem der Ansprüche 17 bis 24.
26. Verfahren nach Anspruch 25 zur Herstellung eines Wertstoffs, insbesondere einer niedermolekularen Verbindung oder eines Proteins.
27. Verfahren nach Anspruch 26, wobei es sich bei der niedermolekularen Verbindung um einen Naturstoff, einen Nahrungsmittelergänzungsstoff oder um eine pharmazeutisch relevante Verbindung handelt.
28. Verfahren nach Anspruch 26, wobei es sich bei dem Protein um ein Enzym handelt, insbesondere eines aus der Gruppe der α -Amylasen, Proteasen, Cellulasen, Lipasen, Oxidoreduktasen, Peroxidasen, Laccasen, Oxidasen und Hemicellulasen.

29. Verwendung des Faktors RecA nach einem der Ansprüche 1 bis 4 und/oder eines RecA, welches mit der in SEQ ID NO. 32 angegebenen Aminosäuresequenz in mindestens 347, vorzugsweise 348 der dort gezeigten 348 Aminosäurepositionen übereinstimmt, in einem molekularbiologischen Reaktionsansatz.
30. Verwendung nach Anspruch 29 zum Stabilisieren einzelsträngiger DNA, insbesondere bei einer DNA-Polymerisation, bei *in vitro* erfolgenden Rekombinationsvorgängen, oder zum Überführen doppelsträngiger DNA in einzelsträngige DNA oder umgekehrt.
31. Vektor, enthaltend eine Nukleinsäure nach einem der Ansprüche 5 bis 7.
32. Vektor nach Anspruch 31, wobei es sich um einen Expressionsvektor handelt.
33. Verfahren zur Herstellung eines Faktors RecA nach einem der Ansprüche 1 bis 4.
34. Verfahren nach Anspruch 33, unter Einsatz einer Nukleinsäure nach einem der Ansprüche 5 bis 7, vorzugsweise eines Expressionsvektors nach Anspruch 32, weiter bevorzugt durch Fermentation eines diese Nukleinsäure beziehungsweise diesen Expressionsvektor enthaltenden Wirts.
35. Verwendung der für einen Faktor RecA codierenden Nukleinsäure zur Expression dieses Faktors.
36. Verwendung nach Anspruch 35, um diesen Faktor selbst herzustellen, insbesondere in einem Verfahren nach Anspruch 34, oder um molekularbiologische Aktivitäten der betreffenden Zellen zu modulieren, insbesondere bei *in vivo* erfolgenden Rekombinationsvorgängen.
37. Verwendung der für einen Faktor RecA codierenden Nukleinsäure nach einem der Ansprüche 5 bis 7 und/oder einer für einen Faktor RecA codierenden Nukleinsäure, deren Nukleotidsequenz zu der in SEQ ID NO. 31 angegebenen Nukleotidsequenz in den Positionen 369 bis 1415 zu mindestens 1045, vorzugsweise mindestens 1046, besonders bevorzugt 1047 dieser 1047 Positionen übereinstimmt, zur Inaktivierung

dieses Faktors oder des Gens *recA* in einem *In-vitro*-Ansatz, insbesondere über Wechselwirkung mit einer zugehörigen Nukleinsäure.

38. Eine für eine Teilsequenz von *recA* codierende oder für eine mit *recA* *in vivo* benachbarte Teilsequenz, vorzugsweise weniger als 1.000 bp, besonders bevorzugt weniger als 500 bp entfernt liegende Nukleinsäure gemäß einer der SEQ ID NO. 25 bis 30.
39. Verwendung mindestens einer, vorzugsweise von mindestens zwei einander entgegenorientierten Nukleinsäuren gemäß SEQ ID NO. 25 bis 30 zur Amplifizierung eines *in vivo* davon umschlossenen DNA-Bereichs.
40. Verwendung nach Anspruch 39 zur Amplifizierung eines *recA*-Gens.
41. Verwendung nach Anspruch 39 oder 40 im Rahmen eines Verfahrens nach einem der Ansprüche 8 bis 16.
42. Verwendung nach einem der Ansprüche 39 bis 41 zur Erzeugung eines grampositiven Bakteriums nach einem der Ansprüche 17 bis 24.
43. Eine für eine Teilsequenz von *spoV* codierende oder für eine mit *spoV* *in vivo* benachbarte Teilsequenz, vorzugsweise weniger als 1.000 bp, besonders bevorzugt weniger als 500 bp entfernt liegende Nukleinsäure gemäß einer der SEQ ID NO. 19 bis 24.
44. Verwendung mindestens einer, vorzugsweise von mindestens zwei einander entgegenorientierten Nukleinsäuren gemäß SEQ ID NO. 25 bis 30 zur Amplifizierung eines *in vivo* davon umschlossenen DNA-Bereichs.
45. Verwendung nach Anspruch 44 zur Amplifizierung eines *spoV*-Gens.
46. Verwendung nach Anspruch 44 oder 45 im Rahmen eines Verfahrens nach einem der Ansprüche 13 bis 16.

47. Verwendung nach einem der Ansprüche 44 bis 46 zur Erzeugung eines grampositiven Bakteriums nach einem der Ansprüche 20 bis 24.

Figur 1

1 50
 1 MSDRQAALDM ALKQIEKQFG KGSIMKLGEQ UEURISUVPS GSLALDAALG
 2 MSDRQAALDM ALKQIEKQFG KGSIMKLGEK TDTRISTVPS GSLALDTALG
 3 MSDRQAALDM ALKQIEKQFG KGSIMKLGEK TDTRISTVPS GSLALDTALG
 4 MSDRQAALDM ALKQIEKQFG KGSIMKLGEK TDTRISTVPS GSLALDTALG

51 100
 1 VGGYPRGRII EVYGPESSGK UUVALHAIAE VQQQGGQAAF IDADUALDPV
 2 IGGYPRGRII EVYGPESSGK TTVALHAIAE VQEKGQAAF IDAEHALDPV
 3 IGGYPRGRII EVYGPESSGK TTVALHAIAE VQQQR.TSAF IDAEHALDPV
 4 IGGYPRGRII EVYGPESSGK TTVALHAIAE VQQQR.TSAF IDAEHALDPV

101 150
 1 YAQKLGVNID ELLLSQPDUG EQALEIAEAL VRSGAVDIVV IDSVAALVPK
 2 YAQKLGVNIE ELLLSQPDTG EQALEIAEAL VRSGAVDIVV VDSVAALVPK
 3 YAQKLGVNIE ELLLSQPDTG EQALEIAEAL VRSGAVDIVV VDSVAALVPK
 4 YAQKLGVNIE ELLLSQPDTG EQALEIAEAL VRSGAVDIVV VDSVAALVPK

151 200
 1 AEIEGDMGDS HVGLQARLMS QALRKLSGAI NKSKUIAIFI NQIREKVGVM
 2 AEIEGDMGDS HVGLQARLMS QALRKLSGAI NKSKTIAIFI NQIREKVGVM
 3 AEIEGDMGDS HVGLQARLMS QALRKLSGAI NKSKTIAIFI NQIREKVGVM
 4 AEIEGDMGDS HVGLQARLMS QALRKLSGAI NKSKTIAIFI NQIREKVGVM

201 250
 1 FGNPEUUPGG RALKFYSSVR LEVRRAEQLK QGNDVMGNKU KIKVVKNKVA
 2 FGNPETTPGG RALKFYSSVR LEVRRAEQLK QGNDVMGNKT RIKVVKNKVA
 3 FGNPETTPGG RALKFYSSVR LEVRRAEQLK QGNDVMGNKT KIKVVKNKVA
 4 FGNPETTPGG RALKFYSSVR LEVRRAEQLK QGNDVMGNKT KIKVVKNKVA

251 300
 1 PPFRUAEVDI MYGEGISKEG EIIDLGUEL D IVQKSGAWYS YQEERLGQGR
 2 PPFRTAEVDI MYGEGISKEG EIIDLGTLD IVQKSGSWYS YEEERLGQGR
 3 PPFRTAEVDI MYGEGISKEG EIIDLGTLD IVQKSGSWYS YEEERLGQGR
 4 PPFRTAEVDI MYGEGISKEG EIIDLGTLD IVQKSGSWYS YEEERLGQGR

301 350
 1 ENAKQFLKEN KDILLMIQE Q IREHYGLDUG GAAPAQEDEA QAQEELEF.S
 2 ENAKQFLKEN KDIMLMIQE Q IREHYGLDNN G...VTEKAE EVQEELEFEE
 3 ENAKQFLKEN KDIMLMIQE Q IREHYGLDNN G...VVQQQAE ETQEELEFEE
 4 ENAKQFLKEN KDIMLMIQE Q IREHYGLDNN G...VVQQQAE ETQEELEFEE

Figur 2 / Teil I

1 50
 1 ATGAGTGATC GTCAGGCAGC CTTAGATATG GCGCTTAAAC AAATAGAAAA
 2 ATGAGTGATC GTCAGGCAGC CTTAGATATG GCTCTTAAAGC AAATAGAAAA
 3 ATGAGTGATC GTCAGGCAGC CTTAGATATG GCTCTTAAAC AAATAGAAAA
 4 ATGAGTGATC GTCAGGCAGC CTTAGATATG GCTCTTAAAC AAATAGAAAA

51 100
 1 GCAGTTGGT AAAGGTTCGA TTATGAAACT CGGCGAACAA ACTGAAACGA
 2 ACAATTCCGC AAAGGTTCCA TCATGAAGCT CGGAGAAAAA ACGGATACAA
 3 ACAGTTCCGC AAAGGTTCCA TTATGAAACT GGGAGAAAAG ACAGATACAA
 4 ACAGTTCCGC AAAGGTTCCA TTATGAAACT GGGAGAAAAG ACAGATACAA

101 150
 1 GAATTTCAAC AGTTCCGAGC GGTTCTTAG CGCTCGATGC GGCTCTTGGGA
 2 GAATTTCAAC GGTGCCGAGC GGTTCCCTTG CACTTGATAC CGCTCTCGGA
 3 GAATTTCTAC TGTACCAAGC GGCTCCCTCG CTCTTGATAC AGCACTGGGA
 4 GAATTTCTAC TGTACCAAGC GGCTCCCTCG CTCTTGATAC AGCACTGGGA

151 200
 1 GTGGGCGGAT ACCCGCGCGG CGGGATTATT GAAGTATAACG GGCCTGAAAG
 2 ATAGGCGGAT ACCCGCGCGG ACGGATTATT GAAGTATAACG GACCTGAAAG
 3 ATTGGCGGAT ATCCTCGCGG ACGGATTATT GAAGTATAACG GTCCCTGAAAG
 4 ATTGGCGGAT ATCCTCGCGG ACGGATTATT GAAGTATAACG GTCCCTGAAAG

201 250
 1 CTCGGTAAA ACGACGGTGG CGCTTCATGC GATTGCCGAA GTTCAGCAGC
 2 CTCAGGTAAA ACGACTGTAG CGCTTCACGC AATCGCTGAG GTTCAGGAAA
 3 CTCAGGTAAA ACAACTGTGG CGCTTCATGC GATTGCTGAA GTTCAGCAGC
 4 CTCAGGTAAA ACAACTGTGG CGCTTCATGC GATTGCTGAA GTTCAGCAGC

251 300
 1 AGGGCGGACA AGCGGCGTTC ATCGACGCCG ACACCGCGCT TGATCCCGTC
 2 AAGGCAGACA GGCAGCATTT ATTGATGCCG AGCATGCTCT TGATCCTGTG
 3 A..GCGGACA AGC.GCGTTT ATCGATGCCG AGCATGCGTT AGATCCGGTA
 4 A..GCGGACA AGC.GCGTTT ATCGATGCCG AGCATGCGTT AGATCCGGTA

301 350
 1 TATGCACAAA AGCTGGCGT CAACATTGAT GAGCTTTGC TGTACAGGCC
 2 TACGCGAAA AGCTCGGTGT CAATATCGAA GAGCTGCTGC TTTCTCAGCC
 3 TACGCGAAA AGCTCGGTGT TAACATCGAA GAGCTTTAC TGTCTCAGCC
 4 TACGCGAAA AGCTCGGTGT TAACATCGAA GAGCTTTAC TGTCTCAGCC

Figur 2 / Teil II

351 400

1 TGATACGGGC GAGCAGGCAG TCGAAATCGC TGAAGCCCTT GTCAGAACCG
 2 GGATACGGGA GAGCAGGCAG TGGAGATTGC TGAAGCGCTG GTGCGAACCG
 3 TGACACAGGC GAGCAGGCAG TTGAAATTGC GGAAGCATTG GTTCGAAGCG
 4 TGACACAGGC GAGCAGGCAG TTGAAATTGC GGAAGCATTG GTTCGAAGCG

401 450

1 GAGCGGTGGA TATCGTTGTC ATCGACTCTG TAGCAGCGCT TGTGCCGAAA
 2 GAGCTGTCGA TATCGTAGTC GTTGACTCTG TTGCGCGCT TGTTCCAAAA
 3 GGGCAGTTGA CATTGTCGTT GTCGACTCTG TAGCCGCTCT CGTTCCGAAA
 4 GGGCAGTTGA CATTGTCGTT GTCGACTCTG TAGCCGCTCT CGTTCCGAAA

451 500

1 GCTGAAATCG AAGGAGATAT GGGGGATTCC CACGTCGGTT TGCAGGCCAG
 2 GCTGAAATTG AAGGTGACAT GGGTGATTCA CACGTCGGTT TACAGGCCAG
 3 GCGGAAATTG AAGGCGACAT GGGAGATTG CATGTCGGTT TACAAGCACG
 4 GCGGAAATTG AAGGCGACAT GGGAGATTG CATGTCGGTT TACAAGCACG

501 550

1 ACTGATGTCT CAGGCGCTTC GCAAGCTTTC CGGAGCGATC AATAAATCGA
 2 TCTCATGTCT CAGGCGCTCC GTAAGCTTTC CGGCGCCATC AATAAATCTA
 3 CTTAATGTCT CAAGCGCTTC GTAAGCTTTC AGGGGCCATT AACAAATCGA
 4 CTTAATGTCT CAAGCGCTTC GTAAGCTTTC AGGGGCCATT AACAAATCGA

551 600

1 AGACCATCGC GATCTTATC AACCAAGATTG GTGAAAAAGT CGGTGTCATG
 2 AAACAATCGC AATCTTATT AACCAAGATTG GTGAAAAAGT CGGC GTTATG
 3 AGACAATCGC GATTTCATT AACCAAAATTG GTGAAAAAGT CGGTGTTATG
 4 AGACAATCGC GATTTCATT AACCAAAATTG GTGAAAAAGT CGGTGTTATG

601 650

1 TTTGGAAATC CTGAGACGAC GCCAGGCCGA AGAGCGCTGA AATTCTACTC
 2 TTCGGAAATC CGGAGACGAC ACCGGGCCGA CGCGCGCTGA AATTCTATTC
 3 TTCGGGAACC CGGAAACAAAC ACCTGGCGGC CGTGC GTTGA AATTCTATTC
 4 TTCGGGAACC CGGAAACAAAC ACCTGGCGGC CGTGC GTTGA AATTCTATTC

651 700

1 TTCTGTCCGC CTTGAAGTGC GCCGCGCAGA GCAGCTGAAA CAAGGCAACG
 2 TTCCGTGCGT CTTGAAGTGC GCCGTGCCGA GCAATTAAAG CAGGGCAACG
 3 TTCCGTGCGT CTTGAAGTGC GCCGTGCTGA ACAGCTGAAA CAAGGCAACG
 4 TTCCGTGCGT CTTGAAGTGC GCCGTGCTGA ACAGCTGAAA CAAGGCAACG

Figur 2 / Teil III

701 750

1 ACGTCATGGG GAACAAGACG AAAATCAAAG TCGTGAAAAA CAAAGTGGCA
 2 ACGTTATGGG GAATAAAACG AGAATTAAAG TCGTAAAAAA CAAAGTCGCT
 3 ACGTAATGGG GAACAAAACG AAAATCAAAG TCGTGAAAAA CAAGGTGGCT
 4 ACGTAATGGG GAACAAAACG AAAATCAAAG TCGTGAAAAA CAAGGTGGCT

751 800

1 CCTCCATTCC GGACAGCCGA AGTGGACATT ATGTACGGGG AAGGAATTTC
 2 CCTCCGTTCC GTACGGCTGA AGTGGACATT ATGTACGGTG AAGGAATCTC
 3 CCGCCGTTCC GTACAGCCGA GGTTGACATT ATGTACGGAG AAGGCATTT
 4 CCGCCGTTCC GTACAGCCGA GGTTGACATT ATGTACGGAG AAGGCATTT

801 850

1 AAAAGAAGGG GAAATCATCG ACCTCGGAAC AGAGCTTGAC ATCGTTCAAA
 2 CAAAGAAGGG GAAATCATCG ACCTTGGAAC TGAACTTGAT ATCGTGCAGA
 3 AAAAGAAGGC GAAATCATTG ATCTAGGAAC TGAACTTGAT ATCGTGCAAA
 4 AAAAGAAGGC GAAATCATTG ATCTAGGAAC TGAACTTGAT ATCGTGCAAA

851 900

1 AGAGCGGTGC ATGGTACTCT TATCAGGAGG AACGCCTTGG ACAAGGCCGT
 2 AAAGCGGCTC GTGGTATTCT TATGAAGAAG AACGCCTCGG ACAGGGCCGT
 3 AAAGCGGTTCA ATGGTACTCT TATGAAGAAG AGCGTCTTGG CCAAGGCCGT
 4 AAAGCGGTTCA ATGGTACTCT TATGAAGAAG AGCGTCTTGG CCAAGGCCGT

901 950

1 GAAAACGCCA AACAGTTCTT GAAAGAAAAC AAGGATATCC TTTTGATGAT
 2 GAAAACGCCA AGCAGTTCTT AAAAGAAAAT AAAGACATCA TGCTGATGAT
 3 GAAAATGCAA ACAATTCTT GAAAGAAAAT AAAGATATCA TGCTGATGAT
 4 GAAAATGCAA ACAATTCTT GAAAGAAAAT AAAGATATCA TGCTGATGAT

951 1000

1 TCAAGAGCAG ATCCGGGAGC ACTACGGTTT GGATACTGGA GGCCTGCTC
 2 TCAAGAACAA ATCCGTGAAC ATTACGGTTT GGACAATAAC GGTGTTAC..
 3 CCAGGAGCAA ATT CGCGAAC ATTACGGCTT GGATAATAAC GGAGTAGTGC
 4 CCAGGAGCAA ATT CGCGAAC ATTACGGCTT GGATAATAAC GGAGTAGTGC

1001 1050

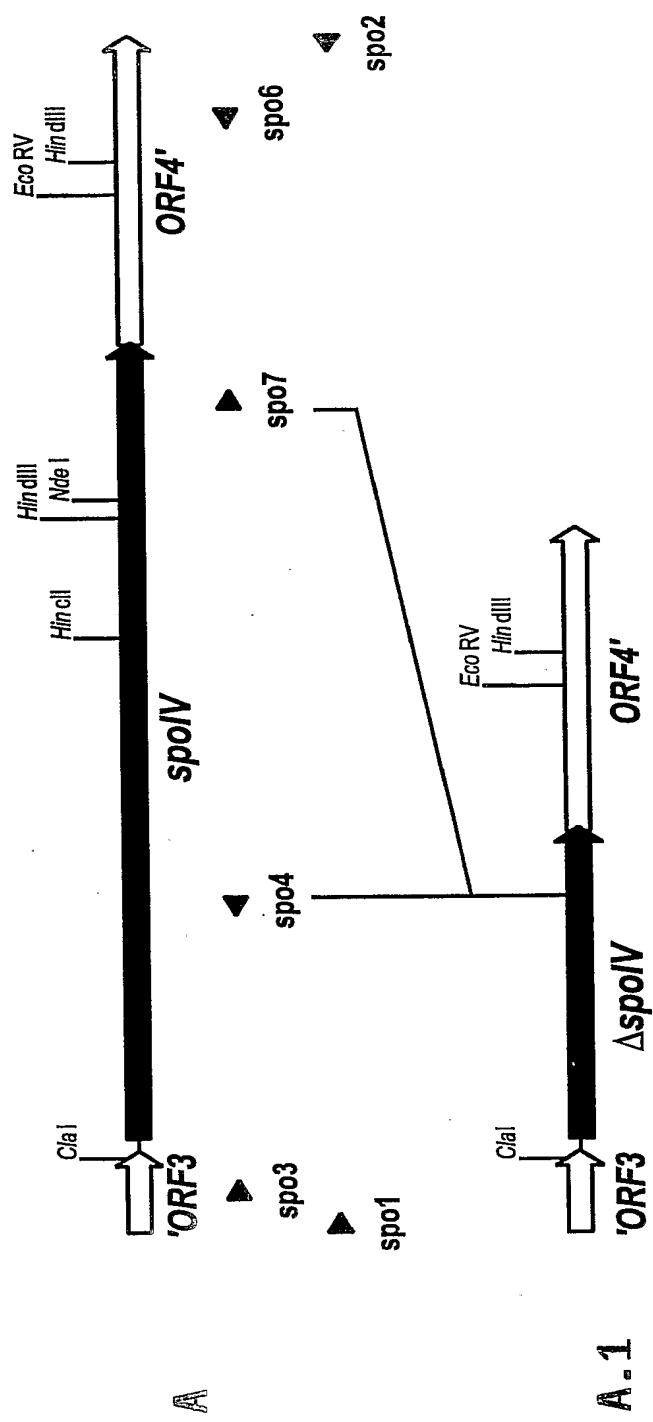
1 CTGCACAGGA AGACGAGGCC CAAGCTCAGG AAGAACTCGA GTTTAATCA
 2GGA AAAAGCGGAA GAAGTTCAGG AAGAGCTTGA ATT CGAAGAA
 3AGCA GCAAGCTGAA GAGACACAAG AAGAACTCGA ATT TGAAAGAA
 4AGCA GCAAGCTGAA GAGACACAAG AAGAACTCGA ATT TGAAAGAA

1051

1 TGA
 2 TAA
 3 ...
 4 TAA

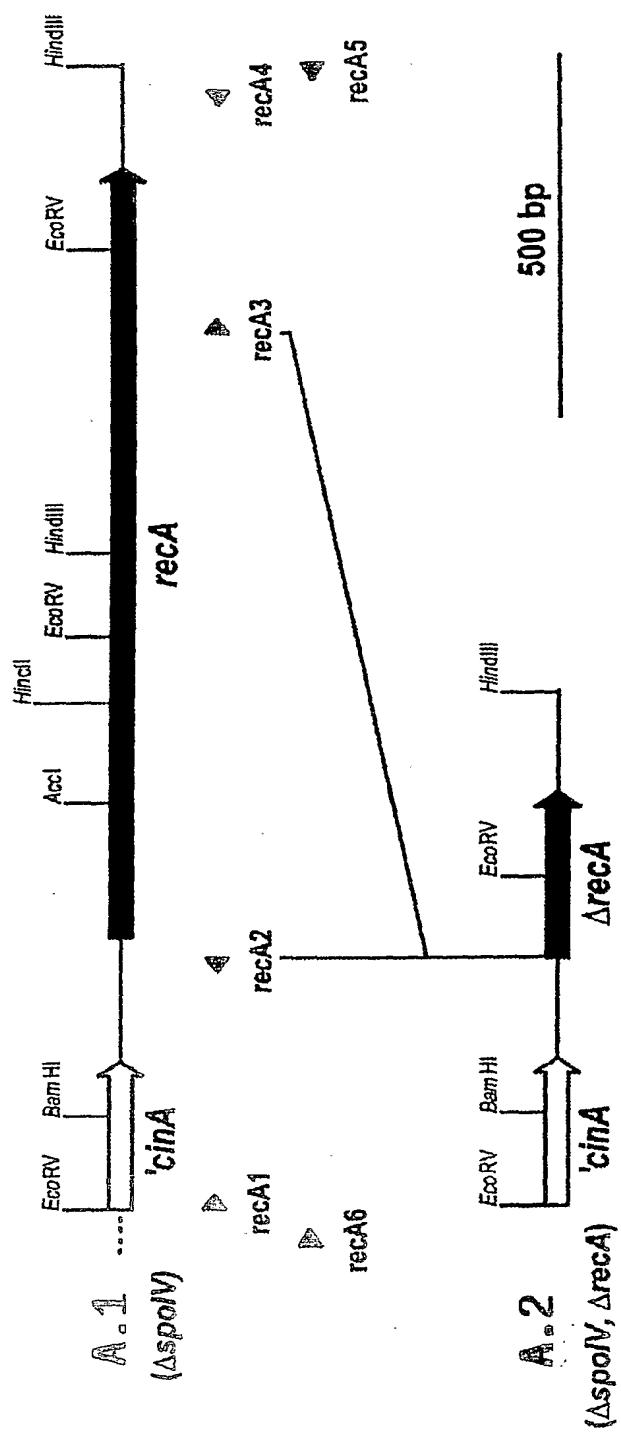
Figur 3

A

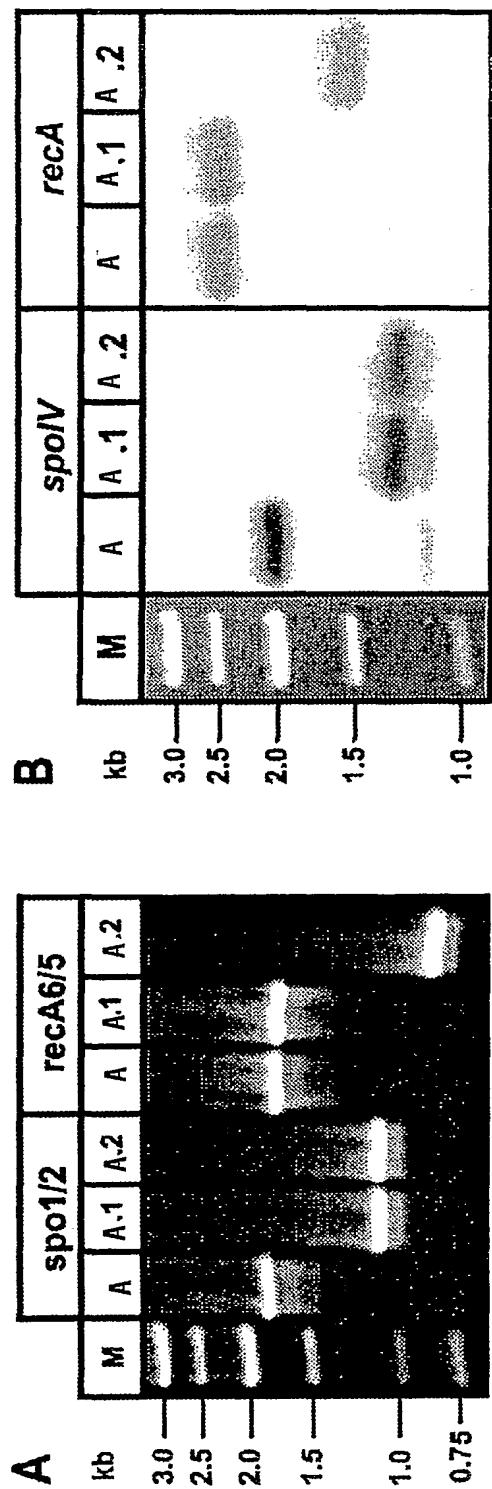


Figur 3

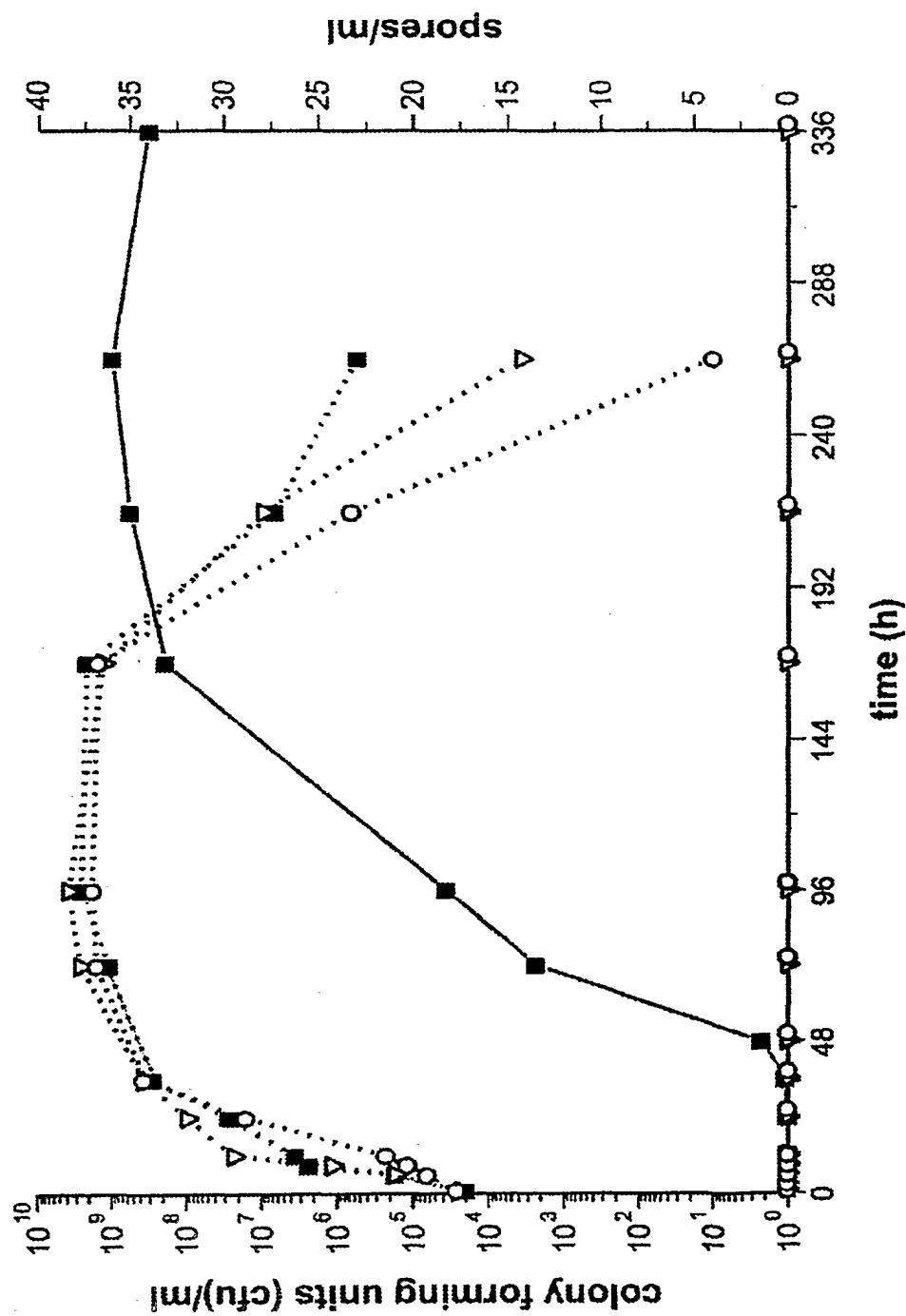
B



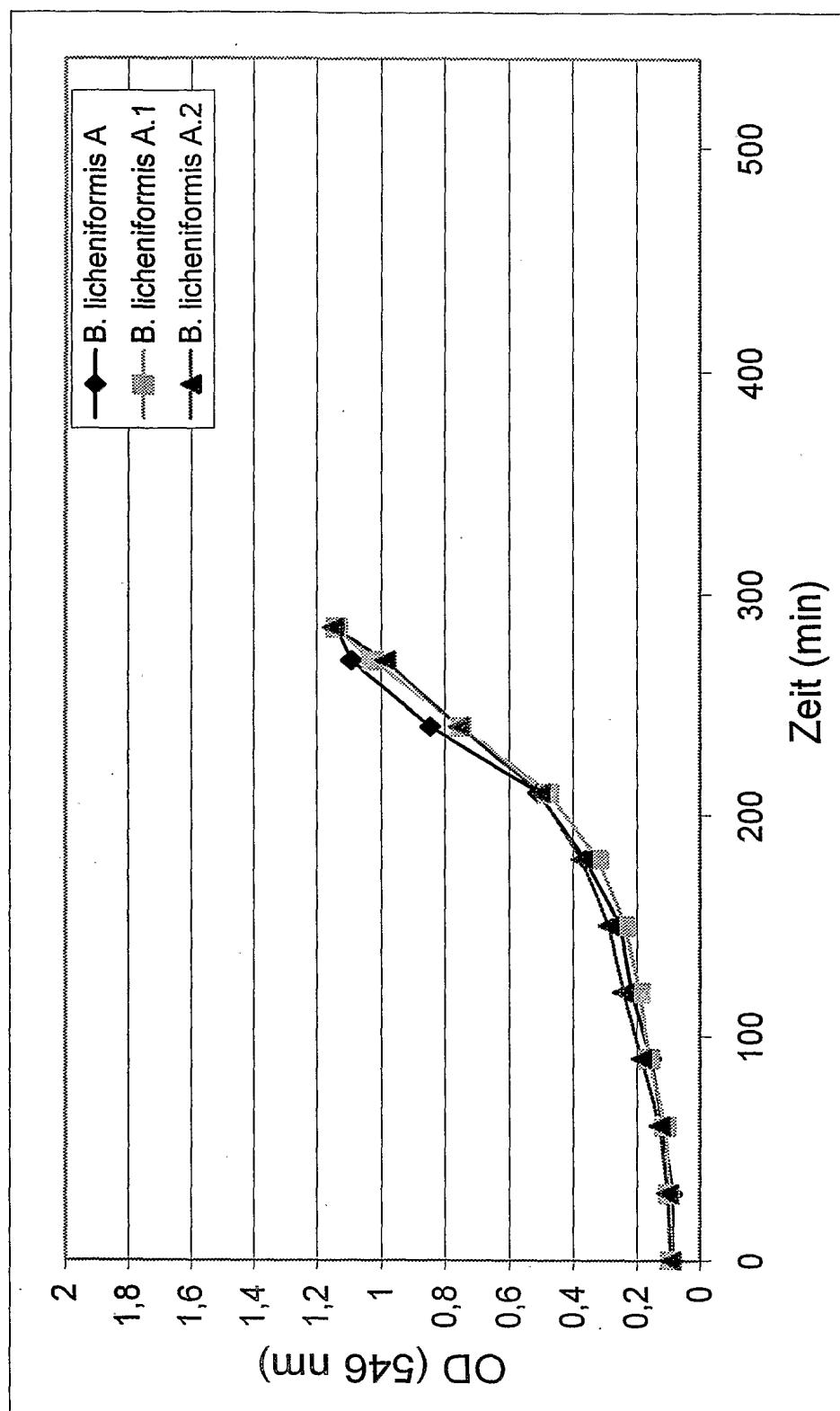
Figur 4



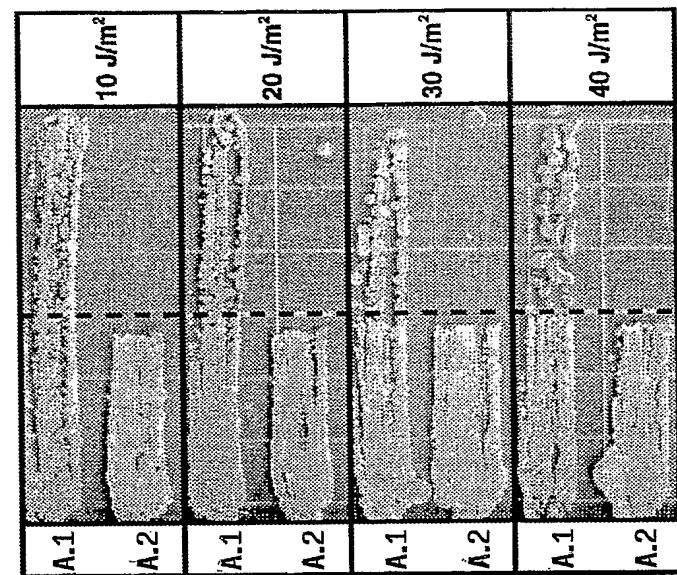
Figur 5



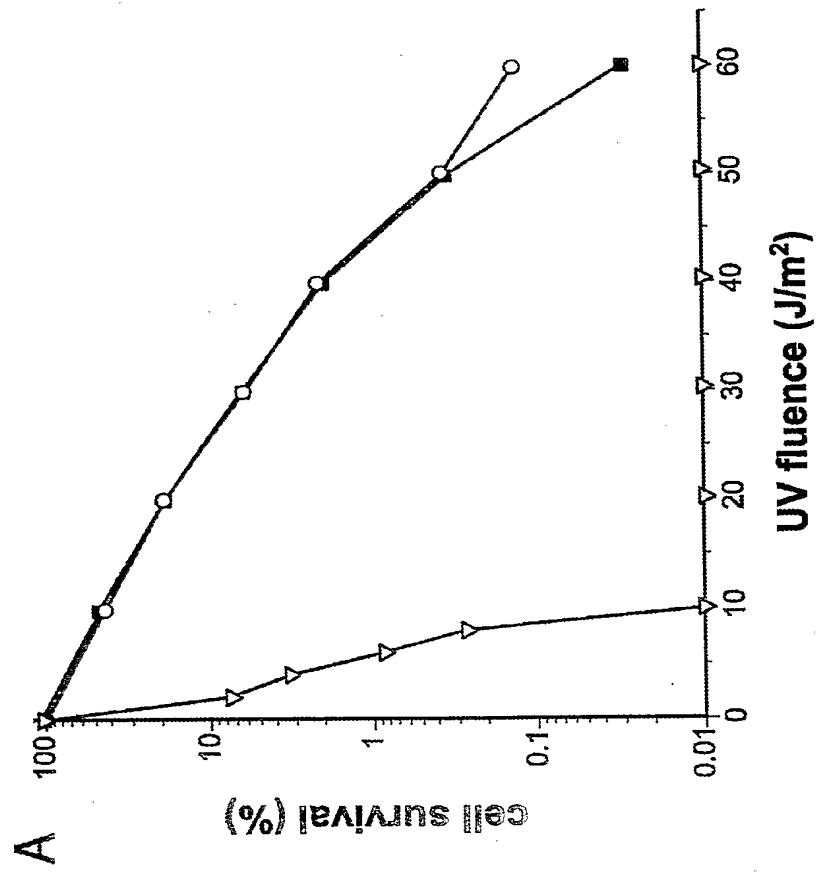
Figur 6



Figur 7



B



SEQUENCE LISTING

<110> Henkel Kommanditgesellschaft auf Aktien

<120> Der Faktor RecA aus *Bacillus licheniformis* und recA-inaktivierte Sicherheitsstämme für die biotechnologische Produktion

<130> H 06291 PCT

<150> DE102004013988

<151> 2004-03-19

<160> 32

<170> PatentIn version 3.1

<210> 1

<211> 1047

<212> DNA

<213> *Bacillus licheniformis* DSM 13

<220>

<221> CDS

<222> (1)..(1047)

<223>

<220>

<221> gene

<222> (1)..(1047)

<223> recA

<400> 1
atg agt gat cgt cag gca gcc tta gat atg gcg ctt aaa caa ata gaa 48
Met Ser Asp Arg Gln Ala Ala Leu Asp Met Ala Leu Lys Gln Ile Glu
1 5 10 15

aag cag ttt ggt aaa ggt tcg att atg aaa ctc ggc gaa caa act gaa 96
Lys Gln Phe Gly Lys Gly Ser Ile Met Lys Leu Gly Glu Gln Thr Glu
20 25 30

acg aga att tca aca gtt ccg agc ggt tct tta gcg ctc gat gcg gct 144
Thr Arg Ile Ser Thr Val Pro Ser Gly Ser Leu Ala Leu Asp Ala Ala
35 40 45

ctt gga gtg ggc gga tac ccg cgc ggc cgg att att gaa gta tac ggg 192
Leu Gly Val Gly Gly Tyr Pro Arg Gly Arg Ile Ile Glu Val Tyr Gly
50 55 60

cct gaa agc tcc ggt aaa acg acg gtg gcg ctt cat gcg att gcc gaa 240
Pro Glu Ser Ser Gly Lys Thr Thr Val Ala Leu His Ala Ile Ala Glu
65 70 75 80

gtt cag cag cag ggc gga caa gcg gcg ttc atc gac gcc gac acc gcg 288
Val Gln Gln Gln Gly Gly Gln Ala Ala Phe Ile Asp Ala Asp Thr Ala
85 90 95

ctt gat ccc gtc tat gca caa aag ctg ggc gtc aac att gat gag ctt 336

Leu	Asp	Pro	Val	Tyr	Ala	Gln	Lys	Leu	Gly	Val	Asn	Ile	Asp	Glu	Leu	
100								105					110			
ttg	ctg	tca	cag	cct	gat	acg	ggc	gag	cag	gcg	ctc	gaa	atc	gct	gaa	384
Leu	Leu	Ser	Gln	Pro	Asp	Thr	Gly	Glu	Gln	Ala	Leu	Glu	Ile	Ala	Glu	
115							120					125				
gcc	ctt	gtc	aga	agc	gga	gcg	gtg	gat	atc	gtt	gtc	atc	gac	tct	gta	432
Ala	Leu	Val	Arg	Ser	Gly	Ala	Val	Asp	Ile	Val	Val	Ile	Asp	Ser	Val	
130							135					140				
gca	gcg	ctt	gtg	ccg	aaa	gct	gaa	atc	gaa	gga	gat	atg	ggg	gat	tcc	480
Ala	Ala	Leu	Val	Pro	Lys	Ala	Glu	Ile	Glu	Gly	Asp	Met	Gly	Asp	Ser	
145							150				155		160			
cac	gtc	ggt	ttg	cag	gcc	aga	ctg	atg	tct	cag	gcg	ctt	cgc	aag	ctt	528
His	Val	Gly	Leu	Gln	Ala	Arg	Leu	Met	Ser	Gln	Ala	Leu	Arg	Lys	Leu	
							165				170		175			
tcc	gga	gcg	atc	aat	aaa	tcg	aag	acc	atc	gcg	atc	ttt	atc	aac	cag	576
Ser	Gly	Ala	Ile	Asn	Lys	Ser	Lys	Thr	Ile	Ala	Ile	Phe	Ile	Asn	Gln	
							180			185		190				
att	cgt	gaa	aaa	gtc	ggt	gtc	atg	ttt	gga	aat	cct	gag	acg	acg	cca	624
Ile	Arg	Glu	Lys	Val	Gly	Val	Met	Phe	Gly	Asn	Pro	Glu	Thr	Thr	Pro	
							195			200		205				
ggc	gga	aga	gcg	ctg	aaa	tcc	tac	tct	tct	gtc	cgc	ctt	gaa	gtg	cgc	672
Gly	Gly	Arg	Ala	Leu	Lys	Phe	Tyr	Ser	Ser	Val	Arg	Leu	Glu	Val	Arg	
							210			215		220				
cgc	gca	gag	cag	ctg	aaa	caa	ggc	aac	gac	gtc	atg	ggg	aac	aag	acg	720
Arg	Ala	Glu	Gln	Leu	Lys	Gln	Gly	Asn	Asp	Val	Met	Gly	Asn	Lys	Thr	
							225			230		235		240		
aaa	atc	aaa	gtc	gtg	aaa	aac	aaa	gtg	gca	cct	cca	ttc	cgg	aca	gcc	768
Lys	Ile	Lys	Val	Val	Lys	Asn	Lys	Val	Ala	Pro	Pro	Phe	Arg	Thr	Ala	
							245			250		255				
gaa	gtg	gac	att	atg	tac	ggg	gaa	gga	att	tca	aaa	gaa	ggg	gaa	atc	816
Glu	Val	Asp	Ile	Met	Tyr	Gly	Glu	Gly	Ile	Ser	Lys	Glu	Gly	Glu	Ile	
							260			265		270				
atc	gac	ctc	gga	aca	gag	ctt	gac	atc	gtt	caa	aag	acg	ggt	gca	tgg	864
Ile	Asp	Leu	Gly	Thr	Glu	Leu	Asp	Ile	Val	Gln	Lys	Ser	Gly	Ala	Trp	
							275			280		285				
tac	tct	tat	cag	gag	gaa	cgc	ctt	gga	caa	ggc	cgt	gaa	aac	gcc	aaa	912
Tyr	Ser	Tyr	Gln	Glu	Arg	Leu	Gly	Gln	Gly	Arg	Glu	Asn	Ala	Lys		
							290			295		300				
cag	ttc	ctg	aaa	gaa	aac	aag	gat	atc	ctt	ttg	atg	att	caa	gag	cag	960
Gln	Phe	Leu	Lys	Glu	Asn	Lys	Asp	Ile	Leu	Leu	Met	Ile	Gln	Glu	Gln	
							305			310		315		320		
atc	cgg	gag	cac	tac	ggt	ttg	gat	act	gga	ggc	gct	gct	cct	gca	cag	1008
Ile	Arg	Glu	His	Tyr	Gly	Leu	Asp	Thr	Gly	Gly	Ala	Ala	Pro	Ala	Gln	
							325			330		335				

gaa gac gag gcc caa gct cag gaa gaa ctc gag ttt taa 1047
Glu Asp Glu Ala Gln Ala Glu Glu Leu Glu Phe
340 345

<210> 2
<211> 348
<212> PRT
<213> *Bacillus licheniformis* DSM 13

<400> 2

Met Ser Asp Arg Gln Ala Ala Leu Asp Met Ala Leu Lys Gln Ile Glu
1 5 10 15

Lys Gln Phe Gly Lys Gly Ser Ile Met Lys Leu Gly Glu Gln Thr Glu
20 25 30

Thr Arg Ile Ser Thr Val Pro Ser Gly Ser Leu Ala Leu Asp Ala Ala
35 40 45

Leu Gly Val Gly Gly Tyr Pro Arg Gly Arg Ile Ile Glu Val Tyr Gly
50 55 60

Pro Glu Ser Ser Gly Lys Thr Thr Val Ala Leu His Ala Ile Ala Glu
65 70 75 80

Val Gln Gln Gln Gly Gln Ala Ala Phe Ile Asp Ala Asp Thr Ala
85 90 95

Leu Asp Pro Val Tyr Ala Gln Lys Leu Gly Val Asn Ile Asp Glu Leu
100 105 110

Leu Leu Ser Gln Pro Asp Thr Gly Glu Gln Ala Leu Glu Ile Ala Glu
115 120 125

Ala Leu Val Arg Ser Gly Ala Val Asp Ile Val Val Ile Asp Ser Val
130 135 140

Ala Ala Leu Val Pro Lys Ala Glu Ile Glu Gly Asp Met Gly Asp Ser
145 150 155 160

His Val Gly Leu Gln Ala Arg Leu Met Ser Gln Ala Leu Arg Lys Leu
165 170 175

Ser Gly Ala Ile Asn Lys Ser Lys Thr Ile Ala Ile Phe Ile Asn Gln
180 185 190

Ile Arg Glu Lys Val Gly Val Met Phe Gly Asn Pro Glu Thr Thr Pro
195 200 205

Gly Gly Arg Ala Leu Lys Phe Tyr Ser Ser Val Arg Leu Glu Val Arg
210 215 220

Arg Ala Glu Gln Leu Lys Gln Gly Asn Asp Val Met Gly Asn Lys Thr
225 230 235 240

Lys Ile Lys Val Val Lys Asn Lys Val Ala Pro Pro Phe Arg Thr Ala
245 250 255

Glu Val Asp Ile Met Tyr Gly Glu Gly Ile Ser Lys Glu Gly Glu Ile
260 265 270

Ile Asp Leu Gly Thr Glu Leu Asp Ile Val Gln Lys Ser Gly Ala Trp
275 280 285

Tyr Ser Tyr Gln Glu Glu Arg Leu Gly Gln Gly Arg Glu Asn Ala Lys
290 295 300

Gln Phe Leu Lys Glu Asn Lys Asp Ile Leu Leu Met Ile Gln Glu Gln
305 310 315 320

Ile Arg Glu His Tyr Gly Leu Asp Thr Gly Gly Ala Ala Pro Ala Gln
325 330 335

Glu Asp Glu Ala Gln Ala Glu Glu Leu Glu Phe
340 345

<210> 3
<211> 1792
<212> DNA
<213> *Bacillus licheniformis*

<220>
<221> CDS
<222> (140)..(1336)
<223>

<220>
<221> gene
<222> (1)..(1792)
<223> spoIV

<220>

```

<221> misc_feature
<222> (140)..(142)
<223> First codon translated as Met.

<400> 3
ggctgatgct caaacagggg cagtgcata ttcaaggcaa agactttgtc atcaaaaacga 60
ttttgcctga gaaaaattctg cttgaaggca cgattgagct tgtccgctat atcgattcat 120
aagtcggggg gaaaagaagc gtg aag aat aaa tgg ctt tct ttt tca gga 172
Val Lys Asn Lys Trp Ile Ser Phe Phe Ser Gly
1 5 10

aag atc cag ctt aag ata acg gga aaa ggg atc gaa cgg tta tta aat 220
Lys Ile Gln Ile Lys Ile Thr Gly Lys Gly Ile Glu Arg Leu Leu Asn
15 20 25

gaa tgc acc agg cgc aac atc ccg atg ttt aat gta aag aaa aag aaa 268
Glu Cys Thr Arg Arg Asn Ile Pro Met Phe Asn Val Lys Lys Lys Lys
30 35 40

gac gcc gtc ttt ctt tat att ccg ctt tct gat gta cat gcc ttc cgg 316
Asp Ala Val Phe Leu Tyr Ile Pro Leu Ser Asp Val His Ala Phe Arg
45 50 55

aag gtc atc aga ggc ttc gac tgc aag tgc agg ttc atc aaa cga aaa 364
Lys Val Ile Arg Gly Phe Asp Cys Lys Cys Arg Phe Ile Lys Arg Lys
60 65 70 75

ggg ttt cct ttc ctc gtg cag aag tct aaa cgg aat agc ggc ttc act 412
Gly Phe Pro Phe Leu Val Gln Lys Ser Lys Arg Asn Ser Gly Phe Thr
80 85 90

ttt gga gtt gct gca ttt ttt atc atc atg ctc cta ttg tcc aac atg 460
Phe Gly Val Ala Ala Phe Phe Ile Ile Met Leu Leu Leu Ser Asn Met
95 100 105

ctt tgg aaa att gat att aca gga gcc aat ccg gag aca gaa cat caa 508
Leu Trp Lys Ile Asp Ile Thr Gly Ala Asn Pro Glu Thr Glu His Gln
110 115 120

atc aaa cag caa ttg gat caa atc ggc gtc aaa aaa ggc cgc ttt cag 556
Ile Lys Gln Gln Leu Asp Gln Ile Gly Val Lys Lys Gly Arg Phe Gln
125 130 135

ttt tca atg ctg acc ccg gaa aaa att cag cag gcg ctc aca aag cgg 604
Phe Ser Met Leu Thr Pro Glu Lys Ile Gln Gln Ala Leu Thr Lys Arg
140 145 150 155

gtc gaa aac atc act tgg gtg ggt att gag tta aac ggc acc gcc ctt 652
Val Glu Asn Ile Thr Trp Val Gly Ile Glu Leu Asn Gly Thr Ala Leu
160 165 170

cac atg aaa gtc gtt gaa aag aat gaa cct gac aaa gaa aaa tat atc 700
His Met Lys Val Val Glu Lys Asn Glu Pro Asp Lys Glu Lys Tyr Ile
175 180 185

ggc ccg agg cac atc gtc gcc aaa aaa ggg gcg acc atc tcg aaa aag 748

```

Gly Pro Arg His Ile Val Ala Lys Lys Gly Ala Thr Ile Ser Lys Lys		
190	195	200
ttc gtg gaa aaa ggc gag ccg ctc gtc acg gtc aac cag cac gtt gaa		796
Phe Val Glu Lys Gly Glu Pro Leu Val Thr Val Asn Gln His Val Glu		
205	210	215
aaa ggg caa atg ctc gtt tcc ggg ctg atc gga agc gaa gag gaa aag		844
Lys Gly Gln Met Leu Val Ser Gly Leu Ile Gly Ser Glu Glu Glu Lys		
220	225	230
caa aaa gtc gga gca aaa ggg aaa atc tac ggt gaa acc tgg tac aag		892
Gln Lys Val Gly Ala Lys Gly Lys Ile Tyr Gly Glu Thr Trp Tyr Lys		
240	245	250
tca aca gta acg gtt cct ctt gag aca tca ttt gac gtt ttt acg ggt		940
Ser Thr Val Thr Val Pro Leu Glu Thr Ser Phe Asp Val Phe Thr Gly		
255	260	265
aaa gta agg aca agt cac aag cta tcc ctc gga tca ttt tcc gtg ccg		988
Lys Val Arg Thr Ser His Lys Leu Ser Leu Gly Ser Phe Ser Val Pro		
270	275	280
atc tgg ggc ttt tca ttt aaa aaa gaa gac ttc tcg cgc ccg aag acg		1036
Ile Trp Gly Phe Ser Phe Lys Lys Glu Asp Phe Ser Arg Pro Lys Thr		
285	290	295
gag acc gaa aac ccc tcg ctg cat ttt atg aat ttt aag ctt cct gtc		1084
Glu Thr Glu Asn Pro Ser Leu His Phe Met Asn Phe Lys Leu Pro Val		
300	305	310
315		
gct tat gaa aag gag cat atg agg gag agc gaa caa atc aaa agg gtg		1132
Ala Tyr Glu Lys Glu His Met Arg Glu Ser Glu Gln Ile Lys Arg Val		
320	325	330
tac tcg aaa aaa gaa gca gtt ctt gaa gga atc gaa atg gga aaa aga		1180
Tyr Ser Lys Glu Ala Val Leu Glu Gly Ile Glu Met Gly Lys Arg		
335	340	345
gac atc agg aaa aaa atc ggc agc gac ggg aac att atc agt gaa aaa		1228
Asp Ile Arg Lys Lys Ile Gly Ser Asp Gly Asn Ile Ile Ser Glu Lys		
350	355	360
gtt ttg cac gaa acg agc gag aat ggc aaa gtt aaa ttg atc atc ctt		1276
Val Leu His Glu Thr Ser Glu Asn Gly Lys Val Lys Leu Ile Ile Leu		
365	370	375
tac cag gtt att gaa gac att gtt caa aca aca cca att gtt cag gag		1324
Tyr Gln Val Ile Glu Asp Ile Val Gln Thr Thr Pro Ile Val Gln Glu		
380	385	390
395		
act aaa gaa tga cagaacactt acttgcaatt catcagcaac tggaaaagtcc		1376
Thr Lys Glu		
gaatgaggct caaacgctgt ttggaaacca ggattcccat ttgaagttga tggaggaaga		1436
gctgaacatt tcaattgtca cgcgcggaga aaccgtgtat gtgacaggag atgaagaaac		1496

gtttgaaatc	gcggacagcc	tgcttccttc	tctcctaaat	ctgatccgca	aaggaatcga	1556
gatataccgaa	cgcgatgtct	tgtatgcgat	caagatggcg	aaaaaggcaga	agcttgagtt	1616
ttttgaaagc	atgtatgaag	aggaaattac	aaaaaacgccc	aaaggaaaac	cgatcagagt	1676
caaaaccatc	ggtcaaagag	aatacatcgc	cggccatgaaa	aggcacgact	taatcttcgg	1736
catcgccca	gcaggaacgg	ggaaaaccta	tttggctgtc	gtaaaggccg	ttcattg	1792

<210> 4
 <211> 398
 <212> PRT
 <213> *Bacillus licheniformis*

<220>
 <221> misc_feature
 <222> (140)..(142)
 <223> First codon translated as Met.

<400> 4

Val	Lys	Asn	Lys	Trp	Leu	Ser	Phe	Phe	Ser	Gly	Lys	Ile	Gln	Leu	Lys
1				5					10					15	

Ile	Thr	Gly	Gly	Ile	Glu	Arg	Leu	Leu	Asn	Glu	Cys	Thr	Arg	Arg
							20		25				30	

Asn	Ile	Pro	Met	Phe	Asn	Val	Lys	Lys	Lys	Asp	Ala	Val	Phe	Leu
			35				40					45		

Tyr	Ile	Pro	Leu	Ser	Asp	Val	His	Ala	Phe	Arg	Lys	Val	Ile	Arg	Gly
						50		55			60				

Phe	Asp	Cys	Lys	Cys	Arg	Phe	Ile	Lys	Arg	Lys	Gly	Phe	Pro	Phe	Leu
			65				70		75				80		

Val	Gln	Lys	Ser	Lys	Arg	Asn	Ser	Gly	Phe	Thr	Phe	Gly	Val	Ala	Ala
								85		90			95		

Phe	Phe	Ile	Ile	Met	Leu	Leu	Leu	Ser	Asn	Met	Leu	Trp	Lys	Ile	Asp
							100			105			110		

Ile	Thr	Gly	Ala	Asn	Pro	Glu	Thr	Glu	His	Gln	Ile	Lys	Gln	Gln	Leu
							115		120			125			

Asp	Gln	Ile	Gly	Val	Lys	Lys	Gly	Arg	Phe	Gln	Phe	Ser	Met	Leu	Thr
							130		135			140			

Pro Glu Lys Ile Gln Gln Ala Leu Thr Lys Arg Val Glu Asn Ile Thr
145 150 155 160

Trp Val Gly Ile Glu Leu Asn Gly Thr Ala Leu His Met Lys Val Val
165 170 175

Glu Lys Asn Glu Pro Asp Lys Glu Lys Tyr Ile Gly Pro Arg His Ile
180 185 190

Val Ala Lys Lys Gly Ala Thr Ile Ser Lys Lys Phe Val Glu Lys Gly
195 200 205

Glu Pro Leu Val Thr Val Asn Gln His Val Glu Lys Gly Gln Met Leu
210 215 220

Val Ser Gly Leu Ile Gly Ser Glu Glu Glu Lys Gln Lys Val Gly Ala
225 230 235 240

Lys Gly Lys Ile Tyr Gly Glu Thr Trp Tyr Lys Ser Thr Val Thr Val
245 250 255

Pro Leu Glu Thr Ser Phe Asp Val Phe Thr Gly Lys Val Arg Thr Ser
260 265 270

His Lys Leu Ser Leu Gly Ser Phe Ser Val Pro Ile Trp Gly Phe Ser
275 280 285

Phe Lys Lys Glu Asp Phe Ser Arg Pro Lys Thr Glu Thr Glu Asn Pro
290 295 300

Ser Leu His Phe Met Asn Phe Lys Leu Pro Val Ala Tyr Glu Lys Glu
305 310 315 320

His Met Arg Glu Ser Glu Gln Ile Lys Arg Val Tyr Ser Lys Lys Glu
325 330 335

Ala Val Leu Glu Gly Ile Glu Met Gly Lys Arg Asp Ile Arg Lys Lys
340 345 350

Ile Gly Ser Asp Gly Asn Ile Ile Ser Glu Lys Val Leu His Glu Thr
355 360 365

Ser Glu Asn Gly Lys Val Lys Leu Ile Ile Leu Tyr Gln Val Ile Glu
370 375 380

Asp Ile Val Gln Thr Thr Pro Ile Val Gln Glu Thr Lys Glu
385 390 395 .

```
<210> 5
<211> 1594
<212> DNA
<213> Bacillus subtilis
```

<220>
<221> CDS
<222> (201)..(1397)
<223>

```
<220>
<221> gene
<222> (1)..(1594)
<223> yqfD
```

```
<220>
<221> misc_feature
<222> (201)..(203)
<223> First codon translated as Met.
```

atc ggt ttt gcg att ttt ttc att ctt ttg ttt ttg ctg tcc aat atg Ile Gly Phe Ala Ile Phe Phe Ile Leu Leu Phe Leu Leu Ser Asn Met 95	100	105	521
gtg tgg aaa att gat gtg aca ggc gct aag cct gaa aca gaa cat caa Val Trp Lys Ile Asp Val Thr Gly Ala Lys Pro Glu Thr Glu His Gln 110	115	120	569
atg agg cag cat ctt aat gaa atc ggc gtc aaa aag ggc cgt ctg cag Met Arg Gln His Leu Asn Glu Ile Gly Val Lys Lys Gly Arg Leu Gln 125	130	135	617
ttt tta atg atg tcg ccc gaa aaa ata cag aaa tca tta acc aat gga Phe Leu Met Met Ser Pro Glu Lys Ile Gln Lys Ser Leu Thr Asn Gly 140	145	150	155
ata gac aat atc act tgg gtc gga gtt gat ctg aag ggg acg acc att Ile Asp Asn Ile Thr Trp Val Gly Val Asp Leu Lys Gly Thr Thr Ile 160	165	170	713
cat atg aaa gtt gtg gag aaa aat gag ccc gaa aaa gaa aaa tat gtt His Met Lys Val Val Glu Lys Asn Glu Pro Glu Lys Glu Lys Tyr Val 175	180	185	761
agc ccg cgc aat att gtc gcc aaa aag aaa gca acc att acg aga atg Ser Pro Arg Asn Ile Val Ala Lys Lys Lys Ala Thr Ile Thr Arg Met 190	195	200	809
tct gtg caa aaa gga cag ccc atg gcc gcc ata cac gat cat gtt gaa Ser Val Gln Lys Gly Gln Pro Met Ala Ala Ile His Asp His Val Glu 205	210	215	857
aag gga cag ctg ctt gtt tcg gga ctg atc ggc agc gaa gac cat cag Lys Gly Gln Leu Leu Val Ser Gly Leu Ile Gly Ser Glu Asp His Gln 220	225	230	235
cag gaa gtc gcc tca aaa gca gaa att tat gga gaa acc tgg tat aga Gln Glu Val Ala Ser Lys Ala Glu Ile Tyr Gly Glu Thr Trp Tyr Arg 240	245	250	953
tca gaa gtg aca gtc ccg ctt gaa aca tta ttt aac gtc tat acg ggc Ser Glu Val Thr Val Pro Leu Glu Thr Leu Phe Asn Val Tyr Thr Gly 255	260	265	1001
aaa gta agg aca aag cac aag ctt tct ttt ggt tct ttg gca atc ccg Lys Val Arg Thr Lys His Lys Leu Ser Phe Gly Ser Leu Ala Ile Pro 270	275	280	1049
atc tgg ggg atg acg ttt aaa aaa gag gaa ttg aag cat cca aaa aca Ile Trp Gly Met Thr Phe Lys Lys Glu Glu Leu Lys His Pro Lys Thr 285	290	295	1097
gaa caa gaa aag cat tcg ctt cat ttt ctc gga ttt aag ctc cct gta Glu Gln Glu Lys His Ser Leu His Phe Leu Gly Phe Lys Leu Pro Val 300	305	310	315
tcc tat gtc aaa gag caa acg aga gaa agt gaa gag gct ttg cga aaa Ser Tyr Val Lys Glu Gln Thr Arg Glu Ser Glu Glu Ala Leu Arg Lys 320	325	330	1193

tat aca aaa gaa gaa gca gtt caa gaa ggc att aaa ttg ggt aaa cag	1241
Tyr Thr Lys Glu Glu Ala Val Gln Glu Gly Ile Lys Leu Gly Lys Gln	
335 340 345	
gat gta gag gat aaa ata ggc gaa aac ggc gag gtg aaa agt gaa aaa	1289
Asp Val Glu Asp Lys Ile Gly Glu Asn Gly Glu Val Lys Ser Glu Lys	
350 355 360	
gtt ttg cac cag act gtt gag aat ggt aaa gta aag ttg att att ctc	1337
Val Leu His Gln Thr Val Glu Asn Gly Lys Val Lys Leu Ile Ile Leu	
365 370 375	
tac caa gtt ata gaa gat atc gtt caa acc aca cct att gtc agg gag	1385
Tyr Gln Val Ile Glu Asp Ile Val Gln Thr Thr Pro Ile Val Arg Glu	
380 385 390 395	
act gaa gaa tga cagaacattt acttgcgatg aatcaaaaac tgaaaaaccc	1437
Thr Glu Glu	
ggacgaggcg ctttcactct tcggaaacca agattcttt ttgaaattga tggagaaaga	1497
tctgaattta aatatcatta cgcgcggcga gacgatttat gtttcaggcg atgatgaatc	1557
gtttcagatt gcagacaggc tgctggatc gtcctc	1594
<210> 6	
<211> 398	
<212> PRT	
<213> <i>Bacillus subtilis</i>	
<220>	
<221> misc_feature	
<222> (201)..(203)	
<223> First codon translated as Met.	
<400> 6	
Val Lys Asn Lys Trp Leu Ser Phe Phe Ser Gly Lys Val Gln Leu Glu	
1 5 10 15	
Leu Thr Gly Arg Gly Ile Glu Arg Leu Leu Asn Glu Cys Thr Lys Gln	
20 25 30	
Gly Ile Pro Val Phe His Val Lys Lys Lys Lys Glu Ala Val Ser Leu	
35 40 45	
Tyr Ile Gln Leu Gln Asp Val His Ala Phe Arg Arg Val Arg Ser Lys	
50 55 60	
Phe Lys Cys Lys Ala Arg Phe Ile Asn Arg Lys Gly Phe Pro Phe Leu	
65 70 75 80	

Leu Leu Lys Ser Lys Leu Asn Ile Gly Phe Thr Ile Gly Phe Ala Ile
85 90 95

Phe Phe Ile Leu Leu Phe Leu Leu Ser Asn Met Val Trp Lys Ile Asp
100 105 110

Val Thr Gly Ala Lys Pro Glu Thr Glu His Gln Met Arg Gln His Leu
115 120 125

Asn Glu Ile Gly Val Lys Lys Gly Arg Leu Gln Phe Leu Met Met Ser
130 135 140

Pro Glu Lys Ile Gln Lys Ser Leu Thr Asn Gly Ile Asp Asn Ile Thr
145 150 155 160

Trp Val Gly Val Asp Leu Lys Gly Thr Thr Ile His Met Lys Val Val
165 170 175

Glu Lys Asn Glu Pro Glu Lys Glu Lys Tyr Val Ser Pro Arg Asn Ile
180 185 190

Val Ala Lys Lys Lys Ala Thr Ile Thr Arg Met Ser Val Gln Lys Gly
195 200 205

Gln Pro Met Ala Ala Ile His Asp His Val Glu Lys Gly Gln Leu Leu
210 215 220

Val Ser Gly Leu Ile Gly Ser Glu Asp His Gln Gln Glu Val Ala Ser
225 230 235 240

Lys Ala Glu Ile Tyr Gly Glu Thr Trp Tyr Arg Ser Glu Val Thr Val
245 250 255

Pro Leu Glu Thr Leu Phe Asn Val Tyr Thr Gly Lys Val Arg Thr Lys
260 265 270

His Lys Leu Ser Phe Gly Ser Leu Ala Ile Pro Ile Trp Gly Met Thr
275 280 285

Phe Lys Lys Glu Glu Leu Lys His Pro Lys Thr Glu Gln Glu Lys His
290 295 300

Ser Leu His Phe Leu Gly Phe Lys Leu Pro Val Ser Tyr Val Lys Glu
305 310 315 320

Gln Thr Arg Glu Ser Glu Glu Ala Leu Arg Lys Tyr Thr Lys Glu Glu
 325 330 335

Ala Val Gln Glu Gly Ile Lys Leu Gly Lys Gln Asp Val Glu Asp Lys
 340 345 350

Ile Gly Glu Asn Gly Glu Val Lys Ser Glu Lys Val Leu His Gln Thr
 355 360 365

Val Glu Asn Gly Lys Val Lys Leu Ile Ile Leu Tyr Gln Val Ile Glu
 370 375 380

Asp Ile Val Gln Thr Thr Pro Ile Val Arg Glu Thr Glu Glu
 385 390 395

<210> 7
 <211> 1876
 <212> DNA
 <213> *Bacillus subtilis*

<220>
 <221> CDS
 <222> (201)..(1679)
 <223>

<220>
 <221> misc_feature
 <222> (201)..(203)
 <223> First codon translated as Met.

<220>
 <221> gene
 <222> (1)..(1876)
 <223> spoIVa

<400> 7
 atgatatgaa aaaggaatga acctttctcc cttgcataca aatagggaga aagggttttt 60
 tatattaata gattgaggat gagaaatttt ctaaagatgt catattcaaa taggacaacg 120
 tcatacacat atagtgtcct gtgttgatt gaaagagctt aataaaattg aaaaggatag 180
 gaagtccggg aggggatcac ttg gaa aag gtc gat att ttc aag gat atc gct 233
 Leu Glu Lys Val Asp Ile Phe Lys Asp Ile Ala
 1 5 10
 gaa cga aca gga ggc gat ata tac tta gga gtc gta ggt gct gtc cgt 281
 Glu Arg Thr Gly Gly Asp Ile Tyr Leu Gly Val Val Gly Ala Val Arg
 15 20 25

aca gga aaa tcc acg ttc att aaa aaa ttt atg gag ctt gtg gtg ctc	329
Thr Gly Lys Ser Thr Phe Ile Lys Lys Phe Met Glu Leu Val Val Leu	
30 35 40	
ccg aat atc agt aac gaa gca gac cgg gcc cga gcg cag gat gaa ctg	377
Pro Asn Ile Ser Asn Glu Ala Asp Arg Ala Arg Ala Gln Asp Glu Leu	
45 50 55	
ccg cag agc gca gcc ggc aaa acc att atg act aca gag cct aaa ttt	425
Pro Gln Ser Ala Ala Gly Lys Thr Ile Met Thr Glu Pro Lys Phe	
60 65 70 75	
gtt ccg aat cag gcg atg tct gtt cat gtg tca gac gga ctc gat gtg	473
Val Pro Asn Gln Ala Met Ser Val His Val Ser Asp Gly Leu Asp Val	
80 85 90	
aat ata aga tta gta gat tgt gta ggt tac aca gtg ccc ggc gct aaa	521
Asn Ile Arg Leu Val Asp Cys Val Gly Tyr Thr Val Pro Gly Ala Lys	
95 100 105	
gga tat gaa gat gaa aac ggg ccg cgg atg atc aat acg cct tgg tac	569
Gly Tyr Glu Asp Glu Asn Gly Pro Arg Met Ile Asn Thr Pro Trp Tyr	
110 115 120	
gaa gaa ccg atc cca ttt cat gag gct gct gaa atc ggc aca cga aaa	617
Glu Glu Pro Ile Pro Phe His Glu Ala Ala Glu Ile Gly Thr Arg Lys	
125 130 135	
gtc att caa gaa cac tcg acc atc gga gtt gtc att acg aca gac ggc	665
Val Ile Gln Glu His Ser Thr Ile Gly Val Val Ile Thr Thr Asp Gly	
140 145 150 155	
acc att gga gat atc gcc aga agt gac tat ata gag gct gaa gaa aga	713
Thr Ile Gly Asp Ile Ala Arg Ser Asp Tyr Ile Glu Ala Glu Glu Arg	
160 165 170	
gtc att gaa gag ctg aaa gag gtt ggc aaa cct ttt att atg gtc atc	761
Val Ile Glu Glu Leu Lys Glu Val Gly Lys Pro Phe Ile Met Val Ile	
175 180 185	
aac tca gtc agg ccg tat cac ccg gaa acg gaa gcc atg cgc cag gat	809
Asn Ser Val Arg Pro Tyr His Pro Glu Thr Glu Ala Met Arg Gln Asp	
190 195 200	
tta agc gaa aaa tat gat atc ccg gta ttg gca atg agt gta gag agc	857
Leu Ser Glu Lys Tyr Asp Ile Pro Val Leu Ala Met Ser Val Glu Ser	
205 210 215	
atg cgg gaa tca gat gtg ctg agt gtg ctc aga gag gcc ctc tac gag	905
Met Arg Glu Ser Asp Val Leu Ser Val Leu Arg Glu Ala Leu Tyr Glu	
220 225 230 235	
ttt ccg gtg cta gaa gtg aat gtc aat ctc cca agc tgg gta atg gtg	953
Phe Pro Val Leu Glu Val Asn Val Asn Leu Pro Ser Trp Val Met Val	
240 245 250	
ctg aaa gaa aac cat tgg ttg cgt gaa agc tat cag gag tcc gtg aag	1001
Leu Lys Glu Asn His Trp Leu Arg Glu Ser Tyr Gln Glu Ser Val Lys	

255	260	265	
gaa acg gtt aag gat att aaa	cg ^g ctc cg ^g gac gta gac agg gtt gtc		1049
Glu Thr Val Lys Asp Ile Lys	Arg Leu Arg Asp Val Asp Arg Val Val		
270	275	280	
ggc caa ttc agc gag ttt gaa	t ^{tc} att gaa agt gcc gga t ^{ta} gcc gga		1097
Gly Gln Phe Ser Glu Phe Glu	Ile Glu Ser Ala Gly Leu Ala Gly		
285	290	295	
att gag ctg ggc caa ggg	gt ^g gca gaa att gat tt ^g tac gc ^g cct gat		1145
Ile Glu Leu Gly Gln Gly Val	Ala Glu Ile Asp Leu Tyr Ala Pro Asp		
300	305	310	315
cat cta tat gat caa atc	cta aaa gaa gtt gt ^g ggc gtc gaa atc aga		1193
His Leu Tyr Asp Gln Ile Leu Lys	Glu Val Val Gly Val Glu Ile Arg		
320	325	330	
gga aga gac cat ctg ctt	gag ctc at ^g caa gac t ^{tc} gcc cat gc ^g aaa		1241
Gly Arg Asp His Leu Leu Glu Leu	Met Gln Asp Phe Ala His Ala Lys		
335	340	345	
aca gaa tat gat caa gt ^g tct	gat gcc t ^{ta} aaa at ^g gtc aaa cag acg		1289
Thr Glu Tyr Asp Gln Val Ser	Asp Ala Leu Lys Met Val Lys Gln Thr		
350	355	360	
gga tac ggc att gca gc ^g	cct gct t ^{ta} gct gat at ^g agt ctc gat gag		1337
Gly Tyr Gly Ile Ala Ala Pro	Ala Leu Ala Asp Met Ser Leu Asp Glu		
365	370	375	
ccg gaa att ata agg cag	ggc tc ^g cga t ^{tc} ggt gt ^g agg ct ^g aaa gct		1385
Pro Glu Ile Ile Arg Gln	Gly Ser Arg Phe Gly Val Arg Leu Lys Ala		
380	385	390	395
gt ^c gct cc ^g tc ^g atc	cat at ^g atc aaa gta gat gt ^c gaa agc gaa t ^{tc}		1433
Val Ala Pro Ser Ile His Met	Ile Lys Val Asp Val Glu Ser Glu Phe		
400	405	410	
gcc cc ^g att atc gga acg	gaa aaa caa agt gaa gag ctt gta cg ^c tat		1481
Ala Pro Ile Ile Gly Thr	Glu Lys Gln Ser Glu Glu Leu Val Arg Tyr		
415	420	425	
tta at ^g cag gac ttt	gag gat cc ^g ctc tcc atc tgg aat tcc gat		1529
Leu Met Gln Asp Phe Glu	Asp Asp Pro Leu Ser Ile Trp Asn Ser Asp		
430	435	440	
atc ttc gga agg tc ^g	ct ^g agc tca att gt ^g aga gaa ggg att cag gca		1577
Ile Phe Gly Arg Ser Leu	Ser Ser Ile Val Arg Glu Gly Ile Gln Ala		
445	450	455	
aag ct ^g tca tt ^g at ^g	cct gaa aac gca cg ^g tat aaa tta aaa gaa aca		1625
Lys Leu Ser Leu Met Pro	Glu Asn Ala Arg Tyr Lys Leu Lys Glu Thr		
460	465	470	475
tta gaa aga atc ata aac	gaa ggc tct ggc ggc t ^{ta} atc gcc atc atc		1673
Leu Glu Arg Ile Ile Asn	Glu Gly Ser Gly Gly Leu Ile Ala Ile Ile		
480	485	490	
ctg taa taccggtaga	ccttttata gaatggagg tctttttct ttgctctaa		1729

Leu

taatggaaaa	ggatcaagga	ataggatgaa	aaaaggaaaa	aaaggaatat	tcgttcggta	1789
aatcacctta	aatccttgac	gagcaaggga	ttgacgctt	aaaatgctt	atatggctt	1849
ttatatgtgt	tactctacat	acagaaa				1876

<210> 8
 <211> 492
 <212> PRT
 <213> *Bacillus subtilis*

<220>
 <221> misc feature
 <222> (201)..(203)
 <223> First codon translated as Met.

<400> 8

Leu	Glu	Lys	Val	Asp	Ile	Phe	Lys	Asp	Ile	Ala	Glu	Arg	Thr	Gly	Gly
1					5				10					15	

Asp	Ile	Tyr	Leu	Gly	Val	Val	Gly	Ala	Val	Arg	Thr	Gly	Lys	Ser	Thr
			20				25						30		

Phe	Ile	Lys	Lys	Phe	Met	Glu	Leu	Val	Val	Leu	Pro	Asn	Ile	Ser	Asn
			35			40						45			

Glu	Ala	Asp	Arg	Ala	Arg	Ala	Gln	Asp	Glu	Leu	Pro	Gln	Ser	Ala	Ala
			50			55				60					

Gly	Lys	Thr	Ile	Met	Thr	Thr	Glu	Pro	Lys	Phe	Val	Pro	Asn	Gln	Ala
65				70			75					80			

Met	Ser	Val	His	Val	Ser	Asp	Gly	Leu	Asp	Val	Asn	Ile	Arg	Leu	Val
					85				90				95		

Asp	Cys	Val	Gly	Tyr	Thr	Val	Pro	Gly	Ala	Lys	Gly	Tyr	Glu	Asp	Glu
			100			105						110			

Asn	Gly	Pro	Arg	Met	Ile	Asn	Thr	Pro	Trp	Tyr	Glu	Glu	Pro	Ile	Pro
					115			120			125				

Phe	His	Glu	Ala	Ala	Glu	Ile	Gly	Thr	Arg	Lys	Val	Ile	Gln	Glu	His
			130			135					140				

Ser	Thr	Ile	Gly	Val	Val	Ile	Thr	Thr	Asp	Gly	Thr	Ile	Gly	Asp	Ile
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

145

150

155

160

Ala Arg Ser Asp Tyr Ile Glu Ala Glu Glu Arg Val Ile Glu Glu Leu
165 170 175

Lys Glu Val Gly Lys Pro Phe Ile Met Val Ile Asn Ser Val Arg Pro
180 185 190

Tyr His Pro Glu Thr Glu Ala Met Arg Gln Asp Leu Ser Glu Lys Tyr
195 200 205

Asp Ile Pro Val Leu Ala Met Ser Val Glu Ser Met Arg Glu Ser Asp
210 215 220

Val Leu Ser Val Leu Arg Glu Ala Leu Tyr Glu Phe Pro Val Leu Glu
225 230 235 240

Val Asn Val Asn Leu Pro Ser Trp Val Met Val Leu Lys Glu Asn His
245 250 255

Trp Leu Arg Glu Ser Tyr Gln Glu Ser Val Lys Glu Thr Val Lys Asp
260 265 270

Ile Lys Arg Leu Arg Asp Val Asp Arg Val Val Gly Gln Phe Ser Glu
275 280 285

Phe Glu Phe Ile Glu Ser Ala Gly Leu Ala Gly Ile Glu Leu Gly Gln
290 295 300

Gly Val Ala Glu Ile Asp Leu Tyr Ala Pro Asp His Leu Tyr Asp Gln
305 310 315 320

Ile Leu Lys Glu Val Val Gly Val Glu Ile Arg Gly Arg Asp His Leu
325 330 335

Leu Glu Leu Met Gln Asp Phe Ala His Ala Lys Thr Glu Tyr Asp Gln
340 345 350

Val Ser Asp Ala Leu Lys Met Val Lys Gln Thr Gly Tyr Gly Ile Ala
355 360 365

Ala Pro Ala Leu Ala Asp Met Ser Leu Asp Glu Pro Glu Ile Ile Arg
370 375 380

Gln Gly Ser Arg Phe Gly Val Arg Leu Lys Ala Val Ala Pro Ser Ile
 385 390 395 400

His Met Ile Lys Val Asp Val Glu Ser Glu Phe Ala Pro Ile Ile Gly
 405 410 415

Thr Glu Lys Gln Ser Glu Glu Leu Val Arg Tyr Leu Met Gln Asp Phe
 420 425 430

Glu Asp Asp Pro Leu Ser Ile Trp Asn Ser Asp Ile Phe Gly Arg Ser
 435 440 445

Leu Ser Ser Ile Val Arg Glu Gly Ile Gln Ala Lys Leu Ser Leu Met
 450 455 460

Pro Glu Asn Ala Arg Tyr Lys Leu Lys Glu Thr Leu Glu Arg Ile Ile
 465 470 475 480

Asn Glu Gly Ser Gly Gly Leu Ile Ala Ile Ile Leu
 485 490

<210> 9
 <211> 1675
 <212> DNA
 <213> *Bacillus subtilis*

<220>
 <221> CDS
 <222> (201)..(1478)
 <223>

<220>
 <221> gene
 <222> (1)..(1675)
 <223> *spoIVB*

<400> 9
 cggatcaagt caaaacaact gggtaagctg cgcgagaagc gcagcttatt ttttcgtgc 60
 acatccatcc gttcatcagt atatccatg ttttcttca tatgacagtt ataaataagc 120
 cgtcagaagg caaaattaaa tggatgtca gcaagtcata aagaagggtgt gggataggag 180
 cgaggagagt gaagtagtga atg ccc gat aac atc aga aaa gca gta ggt tta 233
 Met Pro Asp Asn Ile Arg Lys Ala Val Gly Leu
 1 5 10

att ctc ctt gtt tcg tta tta agt gta ggt tta tgc aaa ccg cta aaa 281
 Ile Leu Leu Val Ser Leu Leu Ser Val Gly Leu Cys Lys Pro Leu Lys
 15 20 25

gaa tat tta ctg att cca acg caa atg aga gta ttt gaa acc caa aca Glu Tyr Leu Leu Ile Pro Thr Gln Met Arg Val Phe Glu Thr Gln Thr 30 35 40	329
caa gcg att gaa acg agt tta tcg gta aat gct cag aca tca gaa tcc Gln Ala Ile Glu Thr Ser Leu Ser Val Asn Ala Gln Thr Ser Glu Ser 45 50 55	377
tca gaa gcg ttt aca gta aag aaa gat ccg cat gaa atc aag gtg acg Ser Glu Ala Phe Thr Val Lys Lys Asp Pro His Glu Ile Lys Val Thr 60 65 70 75	425
ggc aaa aaa tca ggt gag tca gaa ttg gta tat gat ctt gcc gga ttt Gly Lys Ser Gly Glu Ser Glu Leu Val Tyr Asp Leu Ala Gly Phe 80 85 90	473
cca att aaa aaa aca aaa gtg cat gtt ctt cct gat tta aaa gtt ata Pro Ile Lys Lys Thr Lys Val His Val Leu Pro Asp Leu Lys Val Ile 95 100 105	521
cct ggc gga caa tca atc ggt gta aaa ctt cat tcc gtc ggt gtt ctt Pro Gly Gly Gln Ser Ile Gly Val Lys Leu His Ser Val Gly Val Leu 110 115 120	569
gtc gga ttt cat caa atc aat aca agt gaa ggc aaa aaa tct ccg gga Val Gly Phe His Gln Ile Asn Thr Ser Glu Gly Lys Lys Ser Pro Gly 125 130 135	617
gaa acg gca gga att gaa gcg ggc gac atc att att gag atg aat gga Glu Thr Ala Gly Ile Glu Ala Gly Asp Ile Ile Glu Met Asn Gly 140 145 150 155	665
cag aaa att gaa aaa atg aat gat gta gcc cca ttt att caa aag gct Gln Lys Ile Glu Lys Met Asn Asp Val Ala Pro Phe Ile Gln Lys Ala 160 165 170	713
ggg aaa act ggt gaa tct tta gac tta ctg atc aaa cgt gat aaa cag Gly Lys Thr Gly Glu Ser Leu Asp Leu Leu Ile Lys Arg Asp Lys Gln 175 180 185	761
aaa atc aaa acg aag ctg atc cca gaa aag gat gaa gga gaa ggc aaa Lys Ile Lys Thr Lys Leu Ile Pro Glu Lys Asp Glu Gly Glu Gly Lys 190 195 200	809
tac aga atc ggg tta tat atc aga gat tct gct gct ggc atc ggc act Tyr Arg Ile Gly Leu Tyr Ile Arg Asp Ser Ala Ala Gly Ile Gly Thr 205 210 215	857
atg acc ttt tat gaa ccg aaa aca aaa aaa tac gga gca ctt ggc cac Met Thr Phe Tyr Glu Pro Lys Thr Lys Lys Tyr Gly Ala Leu Gly His 220 225 230 235	905
gtg att tcc gat atg gac aca aag aaa cca att gta gtg gag aat gga Val Ile Ser Asp Met Asp Thr Lys Lys Pro Ile Val Val Glu Asn Gly 240 245 250	953
gaa atc gtt aaa tcc act gta aca tca att gaa aaa ggg aca ggc ggt Glu Ile Val Lys Ser Thr Val Thr Ser Ile Glu Lys Gly Thr Gly Gly	1001

	255	260	265	
aat ccg gga gaa aaa ctg gcg cga ttt tcc tca gaa cgc aaa acg atc				1049
Asn Pro Gly Glu Lys Leu Ala Arg Phe Ser Ser Glu Arg Lys Thr Ile				
270	275	280		
ggg gat att aac aga aac agc ccg ttt ggg att ttc ggc aca ctg cat				1097
Gly Asp Ile Asn Arg Asn Ser Pro Phe Gly Ile Phe Gly Thr Leu His				
285	290	295		
cag ccg att caa aac aac ata tca gat caa gca ttg ccg gtt gcg ttt				1145
Gln Pro Ile Gln Asn Asn Ile Ser Asp Gln Ala Leu Pro Val Ala Phe				
300	305	310	315	
tct acc gaa gtc aaa aaa ggg ccg gct gaa att tta acg gtt att gat				1193
Ser Thr Glu Val Lys Lys Gly Pro Ala Glu Ile Leu Thr Val Ile Asp				
320	325	330		
gat gac aaa gta gaa aaa ttc gat att gaa atc gtc agc aca acg ccg				1241
Asp Asp Lys Val Glu Lys Phe Asp Ile Glu Ile Val Ser Thr Thr Pro				
335	340	345		
caa aaa ttc cct ccg aca aaa gga atg gtg ttg aaa att acc gat cca				1289
Gln Lys Phe Pro Ala Thr Lys Gly Met Val Leu Lys Ile Thr Asp Pro				
350	355	360		
aga ctg ttg aaa gaa aca gga ggc atc gta cag ggg atg agc gga agc				1337
Arg Leu Leu Lys Glu Thr Gly Gly Ile Val Gln Gly Met Ser Gly Ser				
365	370	375		
ccg atc att caa aat gga aaa gtg atc ggt gct gtc acc cat gta ttt				1385
Pro Ile Ile Gln Asn Gly Lys Val Ile Gly Ala Val Thr His Val Phe				
380	385	390	395	
gta aat gac ccg aca agc ggc tac ggt gtt cat att gaa tgg atg ctg				1433
Val Asn Asp Pro Thr Ser Gly Tyr Gly Val His Ile Glu Trp Met Leu				
400	405	410		
tca gaa gca gga atc gat att tat gga aaa gaa aaa gca agc tga				1478
Ser Glu Ala Gly Ile Asp Ile Tyr Gly Lys Glu Lys Ala Ser				
415	420	425		
ctgccggagt ttccggcagt ttttttattt tgatccctct tcacttctca gaatacatac				1538
ggtaaaaatat acaaaaagaag atttttcgac aaattcacgt ttcccttgttt gtcaaatttc				1598
attttttagtc gaaaaacaga gaaaaacata gaataacaaa gatatgccac taatattggt				1658
gattatgatt ttttttag				1675

<210> 10

<211> 425

<212> PRT

<213> *Bacillus subtilis*

<400> 10

Met Pro Asp Asn Ile Arg Lys Ala Val Gly Leu Ile Leu Leu Val Ser

21

1

5

10

15

Leu Leu Ser Val Gly Leu Cys Lys Pro Leu Lys Glu Tyr Leu Leu Ile
20 25 30

Pro Thr Gln Met Arg Val Phe Glu Thr Gln Thr Gln Ala Ile Glu Thr
35 40 45

Ser Leu Ser Val Asn Ala Gln Thr Ser Glu Ser Ser Glu Ala Phe Thr
50 55 60

Val Lys Lys Asp Pro His Glu Ile Lys Val Thr Gly Lys Lys Ser Gly
65 70 75 80

Glu Ser Glu Leu Val Tyr Asp Leu Ala Gly Phe Pro Ile Lys Lys Thr
85 90 95

Lys Val His Val Leu Pro Asp Leu Lys Val Ile Pro Gly Gly Gln Ser
100 105 110

Ile Gly Val Lys Leu His Ser Val Gly Val Leu Val Gly Phe His Gln
115 120 125

Ile Asn Thr Ser Glu Gly Lys Lys Ser Pro Gly Glu Thr Ala Gly Ile
130 135 140

Glu Ala Gly Asp Ile Ile Glu Met Asn Gly Gln Lys Ile Glu Lys
145 150 155 160

Met Asn Asp Val Ala Pro Phe Ile Gln Lys Ala Gly Lys Thr Gly Glu
165 170 175

Ser Leu Asp Leu Leu Ile Lys Arg Asp Lys Gln Lys Ile Lys Thr Lys
180 185 190

Leu Ile Pro Glu Lys Asp Glu Gly Glu Gly Lys Tyr Arg Ile Gly Leu
195 200 205

Tyr Ile Arg Asp Ser Ala Ala Gly Ile Gly Thr Met Thr Phe Tyr Glu
210 215 220

Pro Lys Thr Lys Lys Tyr Gly Ala Leu Gly His Val Ile Ser Asp Met
225 230 235 240

Asp Thr Lys Lys Pro Ile Val Val Glu Asn Gly Glu Ile Val Lys Ser
245 250 255

Thr Val Thr Ser Ile Glu Lys Gly Thr Gly Gly Asn Pro Gly Glu Lys
260 265 270

Leu Ala Arg Phe Ser Ser Glu Arg Lys Thr Ile Gly Asp Ile Asn Arg
275 280 285

Asn Ser Pro Phe Gly Ile Phe Gly Thr Leu His Gln Pro Ile Gln Asn
290 295 300

Asn Ile Ser Asp Gln Ala Leu Pro Val Ala Phe Ser Thr Glu Val Lys
305 310 315 320

Lys Gly Pro Ala Glu Ile Leu Thr Val Ile Asp Asp Asp Lys Val Glu
325 330 335

Lys Phe Asp Ile Glu Ile Val Ser Thr Thr Pro Gln Lys Phe Pro Ala
340 345 350

Thr Lys Gly Met Val Leu Lys Ile Thr Asp Pro Arg Leu Leu Lys Glu
355 360 365

Thr Gly Gly Ile Val Gln Gly Met Ser Gly Ser Pro Ile Ile Gln Asn
370 375 380

Gly Lys Val Ile Gly Ala Val Thr His Val Phe Val Asn Asp Pro Thr
385 390 395 400

Ser Gly Tyr Gly Val His Ile Glu Trp Met Leu Ser Glu Ala Gly Ile
405 410 415

Asp Ile Tyr Gly Lys Glu Lys Ala Ser
420 425

<210> 11
<211> 1900
<212> DNA
<213> *Bacillus subtilis*

<220>
<221> CDS
<222> (201)..(1703)
<223>

<220>
 <221> gene
 <222> (1)..(1900)
 <223> spoIVCA

<220>
 <221> misc_feature
 <222> (201)..(203)
 <223> First codon translated as Met.

<400> 11
 ttttgcata tcattgaaac gtttaataac actatagttt aatttaaaat ctcctcattt 60
 ggacaaacag ctgttacata gcattaccga aggggtgatg cattttatga aagtgataat 120
 catcgaggga ccgcaagctg acaaatgcat taacgattgc tatttattttaataaaaact 180
 ttataggaag gagattcagg gtg ata gca ata tat gta agg gta tcg acc gag 233
 Val Ile Ala Ile Tyr Val Arg Val Ser Thr Glu
 1 5 10
 gaa caa gcg atc aag gga tcg agc atc gac agc caa atc gag gcc tgt 281
 Glu Gln Ala Ile Lys Gly Ser Ser Ile Asp Ser Gln Ile Glu Ala Cys
 15 20 25
 ata aag aaa gca ggg act aaa gat gtg ctg aag tat gca gat gaa gga 329
 Ile Lys Ala Gly Thr Lys Asp Val Leu Lys Tyr Ala Asp Glu Gly
 30 35 40
 ttt tca gga gag ctt tta gaa cgt ccg gct ttg aat cgc ttg agg gag 377
 Phe Ser Gly Glu Leu Leu Glu Arg Pro Ala Leu Asn Arg Leu Arg Glu
 45 50 55
 gat gca agc aag gga ctt ata agt caa gtc att tgt tac gat cct gac 425
 Asp Ala Ser Lys Gly Leu Ile Ser Gln Val Ile Cys Tyr Asp Pro Asp
 60 65 70 75
 cgt ctt tct cgg aaa tta atg aat cag cta atc att gat gac gaa ttg 473
 Arg Leu Ser Arg Lys Leu Met Asn Gln Leu Ile Ile Asp Asp Glu Leu
 80 85 90
 cga aag cga aac ata cct ttg att ttt gta aat ggt gaa tac gcc aat 521
 Arg Lys Arg Asn Ile Pro Leu Ile Phe Val Asn Gly Glu Tyr Ala Asn
 95 100 105
 tct cca gaa ggt caa ttg ttt ttc gca atg cgc ggg gca atc tca gaa 569
 Ser Pro Glu Gly Gln Leu Phe Phe Ala Met Arg Gly Ala Ile Ser Glu
 110 115 120
 ttt gaa aaa gcc aaa atc aaa gaa cgg aca tca agc ggc cga ctt caa 617
 Phe Glu Lys Ala Lys Ile Lys Glu Arg Thr Ser Ser Gly Arg Leu Gln
 125 130 135
 aaa atg aaa aaa ggc atg atc att aaa gat tct aaa cta tat ggc tat 665
 Lys Met Lys Lys Gly Met Ile Ile Lys Asp Ser Lys Leu Tyr Gly Tyr
 140 145 150 155

aaa ttt gtt aaa gag aaa aga act ctt gag ata tta gaa gag gaa gca	713
Lys Phe Val Lys Glu Lys Arg Thr Leu Glu Ile Leu Glu Glu Ala	
160 165 170	
aaa atc att cggtt att ttt aac tat ttc acc gat cat aaa agc cct	761
Lys Ile Ile Arg Met Ile Phe Asn Tyr Phe Thr Asp His Lys Ser Pro	
175 180 185	
ttt ttc ggc aga gta aat ggt att gct cta cat tta act cag atg ggg	809
Phe Phe Gly Arg Val Asn Gly Ile Ala Leu His Leu Thr Gln Met Gly	
190 195 200	
gtt aaa aca aaa aaa ggc gcc aaa gta tgg cac agg cag gtt gtt cggtt	857
Val Lys Thr Lys Lys Gly Ala Lys Val Trp His Arg Gln Val Val Arg	
205 210 215	
caa ata tta atg aac tct tcc tat aag ggt gaa cat aga cag tat aaa	905
Gln Ile Leu Met Asn Ser Ser Tyr Lys Gly Glu His Arg Gln Tyr Lys	
220 225 230 235	
tat gat aca gag ggt tcc tat gtt tca aag cag gca ggg aac aaa tct	953
Tyr Asp Thr Glu Gly Ser Tyr Val Ser Lys Gln Ala Gly Asn Lys Ser	
240 245 250	
ata att aaa ata agg cct gaa gaa gaa caa atc act gtg aca att cca	1001
Ile Ile Lys Ile Arg Pro Glu Glu Gln Ile Thr Val Thr Ile Pro	
255 260 265	
gca att gtt cca gct gaa caa tgg gat tat gct caa gaa ctc tta ggt	1049
Ala Ile Val Pro Ala Glu Gln Trp Asp Tyr Ala Gln Glu Leu Leu Gly	
270 275 280	
caa agt aaa aga aaa cac ttg agt atc agc cct cac aat tac ttg tta	1097
Gln Ser Lys Arg Lys His Leu Ser Ile Ser Pro His Asn Tyr Leu Leu	
285 290 295	
tcg ggt ttg gtt aga tgc gga aaa tgc gga aat acc atg aca ggg aag	1145
Ser Gly Leu Val Arg Cys Gly Lys Cys Gly Asn Thr Met Thr Gly Lys	
300 305 310 315	
aaa aga aaa tca cat ggt aaa gac tac tat gta tat act tgc cgg aaa	1193
Lys Arg Lys Ser His Gly Lys Asp Tyr Tyr Val Tyr Thr Cys Arg Lys	
320 325 330	
aat tat tct ggc gca aag gac cgc ggc tgc gga aaa gaa atg tct gag	1241
Asn Tyr Ser Gly Ala Lys Asp Arg Gly Cys Gly Lys Glu Met Ser Glu	
335 340 345	
aat aaa ttg aac cgg cat gta tgg ggt gaa att ttt aaa ttc atc aca	1289
Asn Lys Leu Asn Arg His Val Trp Gly Glu Ile Phe Lys Phe Ile Thr	
350 355 360	
aat cct caa aag tat gtt tct ttt aaa gag gct gaa caa tca aat cac	1337
Asn Pro Gln Lys Tyr Val Ser Phe Lys Glu Ala Glu Gln Ser Asn His	
365 370 375	
ctg tct gat gaa tta gaa ctt att gaa aaa gag ata gag aaa aca aaa	1385
Leu Ser Asp Glu Leu Glu Leu Ile Glu Lys Glu Ile Glu Lys Thr Lys	
380 385 390 395	

aaa ggc cgc aag cgt ctt tta acg cta atc agc cta agc gat gac gat	400	405	410	1433
Lys Gly Arg Lys Arg Leu Leu Thr Leu Ile Ser Leu Ser Asp Asp Asp				
gat tta gac ata gat gaa atc aaa gca caa att att gaa ctg caa aaa	415	420	425	1481
Asp Leu Asp Ile Asp Glu Ile Lys Ala Gln Ile Ile Glu Leu Gln Lys				
aag caa aat cag ctt act gaa aag tgt aac aga atc cag tca aaa atg	430	435	440	1529
Lys Gln Asn Gln Leu Thr Glu Lys Cys Asn Arg Ile Gln Ser Lys Met				
aaa gtc cta gat gat acg agc tca agt gaa aat gct cta aaa aga gcc	445	450	455	1577
Lys Val Leu Asp Asp Thr Ser Ser Glu Asn Ala Leu Lys Arg Ala				
atc gac tat ttt caa tca atc ggt gca gat aac tta act ctt gaa gat	460	465	470	1625
Ile Asp Tyr Phe Gln Ser Ile Gly Ala Asp Asn Leu Thr Leu Glu Asp				
aaa aaa aca att gtt aac ttt atc gtg aaa gaa gtt acc att gtg gat	480	485	490	1673
Lys Lys Thr Ile Val Asn Phe Ile Val Lys Glu Val Thr Ile Val Asp				
tct gac acc ata tat att gaa acg tat taa agaggggtgt atgcacccccc	495	500		1723
Ser Asp Thr Ile Tyr Ile Glu Thr Tyr				
cttttgtaat tacaatctca ttttcaatac acctcgctgc atacgtcgcc acctttgtcc				1783
cttttccagc ggaatagctt tcaattcctt taataagccc gatcggtccg atggagatta				1843
agtccctctgc atcctcacct gtatttcga acttttcac aatatggcgc accaagc				1900
<210> 12				
<211> 500				
<212> PRT				
<213> Bacillus subtilis				
<220>				
<221> misc_feature				
<222> (201)..(203)				
<223> First codon translated as Met.				
<400> 12				
Val Ile Ala Ile Tyr Val Arg Val Ser Thr Glu Glu Gln Ala Ile Lys				
1	5	10	15	
Gly Ser Ser Ile Asp Ser Gln Ile Glu Ala Cys Ile Lys Lys Ala Gly				
20	25	30		
Thr Lys Asp Val Leu Lys Tyr Ala Asp Glu Gly Phe Ser Gly Glu Leu				
35	40	45		

Leu Glu Arg Pro Ala Leu Asn Arg Leu Arg Glu Asp Ala Ser Lys Gly
50 55 60

Leu Ile Ser Gln Val Ile Cys Tyr Asp Pro Asp Arg Leu Ser Arg Lys
65 70 75 80

Leu Met Asn Gln Leu Ile Ile Asp Asp Glu Leu Arg Lys Arg Asn Ile
85 90 95

Pro Leu Ile Phe Val Asn Gly Glu Tyr Ala Asn Ser Pro Glu Gly Gln
100 105 110

Leu Phe Phe Ala Met Arg Gly Ala Ile Ser Glu Phe Glu Lys Ala Lys
115 120 125

Ile Lys Glu Arg Thr Ser Ser Gly Arg Leu Gln Lys Met Lys Lys Gly
130 135 140

Met Ile Ile Lys Asp Ser Lys Leu Tyr Gly Tyr Lys Phe Val Lys Glu
145 150 155 160

Lys Arg Thr Leu Glu Ile Leu Glu Glu Ala Lys Ile Ile Arg Met
165 170 175

Ile Phe Asn Tyr Phe Thr Asp His Lys Ser Pro Phe Phe Gly Arg Val
180 185 190

Asn Gly Ile Ala Leu His Leu Thr Gln Met Gly Val Lys Thr Lys Lys
195 200 205

Gly Ala Lys Val Trp His Arg Gln Val Val Arg Gln Ile Leu Met Asn
210 215 220

Ser Ser Tyr Lys Gly Glu His Arg Gln Tyr Lys Tyr Asp Thr Glu Gly
225 230 235 240

Ser Tyr Val Ser Lys Gln Ala Gly Asn Lys Ser Ile Ile Lys Ile Arg
245 250 255

Pro Glu Glu Glu Gln Ile Thr Val Thr Ile Pro Ala Ile Val Pro Ala
260 265 270

Glu Gln Trp Asp Tyr Ala Gln Glu Leu Leu Gly Gln Ser Lys Arg Lys
275 280 285

His Leu Ser Ile Ser Pro His Asn Tyr Leu Leu Ser Gly Leu Val Arg
290 295 300

Cys Gly Lys Cys Gly Asn Thr Met Thr Gly Lys Lys Arg Lys Ser His
305 310 315 320

Gly Lys Asp Tyr Tyr Val Tyr Thr Cys Arg Lys Asn Tyr Ser Gly Ala
325 330 335

Lys Asp Arg Gly Cys Gly Lys Glu Met Ser Glu Asn Lys Leu Asn Arg
340 345 350

His Val Trp Gly Glu Ile Phe Lys Phe Ile Thr Asn Pro Gln Lys Tyr
355 360 365

Val Ser Phe Lys Glu Ala Glu Gln Ser Asn His Leu Ser Asp Glu Leu
370 375 380

Glu Leu Ile Glu Lys Glu Ile Glu Lys Thr Lys Lys Gly Arg Lys Arg
385 390 395 400

Leu Leu Thr Leu Ile Ser Leu Ser Asp Asp Asp Asp Leu Asp Ile Asp
405 410 415

Glu Ile Lys Ala Gln Ile Ile Glu Leu Gln Lys Lys Gln Asn Gln Leu
420 425 430

Thr Glu Lys Cys Asn Arg Ile Gln Ser Lys Met Lys Val Leu Asp Asp
435 440 445

Thr Ser Ser Ser Glu Asn Ala Leu Lys Arg Ala Ile Asp Tyr Phe Gln
450 455 460

Ser Ile Gly Ala Asp Asn Leu Thr Leu Glu Asp Lys Lys Thr Ile Val
465 470 475 480

Asn Phe Ile Val Lys Glu Val Thr Ile Val Asp Ser Asp Thr Ile Tyr
485 490 495

Ile Glu Thr Tyr
500

<211> 868
 <212> DNA
 <213> *Bacillus subtilis*

<220>
 <221> CDS
 <222> (201)..(671)
 <223>

<220>
 <221> gene
 <222> (1)..(868)
 <223> *spoIVCB*

<400> 13
 ttcatccca tccccccata cctttgttca tttcaatgta tggcgcttg atgaagaata 60
 ttttaacat ttgaagttag tatgctgctt accaaagccg gactcccccg cgagaaattt 120
 cccggtagac acacagacag cctcccggtc acatacattt acatataggc tttgcctac 180
 atactttgt ggaggtgacg atg gtg aca ggt ttc gca gcg ctc ggc ttt 233
 Met Val Thr Gly Val Phe Ala Ala Leu Gly Phe
 1 5 10

gtt gtt aaa gag ctt gtc ttt tta gta tct tac gtg aaa aac aat gcc 281
 Val Val Lys Glu Leu Val Phe Leu Val Ser Tyr Val Lys Asn Asn Ala
 15 20 25

ttt cca caa ccg ctc tca agc agc gaa gaa aaa aaa tac tta gag ctc 329
 Phe Pro Gln Pro Leu Ser Ser Glu Glu Lys Lys Tyr Leu Glu Leu
 30 35 40

atg gct aaa ggg gat gaa cat gcc aga aac atg ctg att gag cat aat 377
 Met Ala Lys Gly Asp Glu His Ala Arg Asn Met Leu Ile Glu His Asn
 45 50 55

ctt cgc ttg gtc gcc cat att gtg aaa aag ttc gaa aat aca ggt gag 425
 Leu Arg Leu Val Ala His Ile Val Lys Lys Phe Glu Asn Thr Gly Glu
 60 65 70 75

gat gca gag gac tta atc tcc atc gga acg atc ggg ctt att aaa gga 473
 Asp Ala Glu Asp Leu Ile Ser Ile Gly Thr Ile Gly Leu Ile Lys Gly
 80 85 90

att gaa agc tat tcc gct gga aaa ggg aca aag gtg gcg acg tat gca 521
 Ile Glu Ser Tyr Ser Ala Gly Lys Gly Thr Lys Val Ala Thr Tyr Ala
 95 100 105

gcg agg tgt att gaa aat gag att gta att aca aaa ggg ggg tgc ata 569
 Ala Arg Cys Ile Glu Asn Glu Ile Val Ile Thr Lys Gly Gly Cys Ile
 110 115 120

cac ccc tct tta ata cgt ttc aat ata tat ggt gtc aga atc cac aat 617
 His Pro Ser Leu Ile Arg Phe Asn Ile Tyr Gly Val Arg Ile His Asn
 125 130 135

ggt aac ttc ttt cac gat aaa gtt aac aat tgt ttt ttt atc ttc aag 665
 Gly Asn Phe Phe His Asp Lys Val Asn Asn Cys Phe Phe Ile Phe Lys
 140 145 150 155

agt taa gttatctgca ccgattgatt gaaaatagtc gatggctctt ttttagagcat 721
 Ser

tttcacttga gctcgtatca tctaggactt tcattttga ctggattctg ttacacttt 781
 cagtaagctg attttgctt ttttcagtt caataatgg tgcttgatt tcatctatgt 841
 ctaaatcatc gtcatcgctt aggctga 868

<210> 14
 <211> 156
 <212> PRT
 <213> *Bacillus subtilis*

<400> 14

Met Val Thr Gly Val Phe Ala Ala Leu Gly Phe Val Val Lys Glu Leu
 1 5 10 15

Val Phe Leu Val Ser Tyr Val Lys Asn Asn Ala Phe Pro Gln Pro Leu
 20 25 30

Ser Ser Ser Glu Glu Lys Lys Tyr Leu Glu Leu Met Ala Lys Gly Asp
 35 40 45

Glu His Ala Arg Asn Met Leu Ile Glu His Asn Leu Arg Leu Val Ala
 50 55 60

His Ile Val Lys Lys Phe Glu Asn Thr Gly Glu Asp Ala Glu Asp Leu
 65 70 75 80

Ile Ser Ile Gly Thr Ile Gly Leu Ile Lys Gly Ile Glu Ser Tyr Ser
 85 90 95

Ala Gly Lys Gly Thr Lys Val Ala Thr Tyr Ala Ala Arg Cys Ile Glu
 100 105 110

Asn Glu Ile Val Ile Thr Lys Gly Gly Cys Ile His Pro Ser Leu Ile
 115 120 125

Arg Phe Asn Ile Tyr Gly Val Arg Ile His Asn Gly Asn Phe Phe His
 130 135 140

Asp Lys Val Asn Asn Cys Phe Phe Ile Phe Lys Ser

145	150	155	
<210> 15			
<211> 1192			
<212> DNA			
<213> <i>Bacillus subtilis</i>			
<220>			
<221> CDS			
<222> (201)..(995)			
<223>			
<220>			
<221> gene			
<222> (1)..(1192)			
<223> spoIVFA			
<400> 15			
acaaaggaat gatggctaag attaagtcat ttttcggagt aagatcttaa tgtgatagaa 60			
tcaaagagaa gaatctgaca aagcatatgc tgtgtcaggt tttttttgt ttttgccgc 120			
tttgttcttg actaaaccga atatggcca tggacaagac atatgatgta caaacccaa 180			
gaatgcaaag gatgatggca atg agt cac aga gca gat gaa atc aga aaa cga 233			
Met Ser His Arg Ala Asp Glu Ile Arg Lys Arg			
1 5 10			
tta gag aaa aga aga aag cag ctt tcc ggc tca aaa cgt ttc tct act 281			
Leu Glu Lys Arg Arg Lys Gln Leu Ser Gly Ser Lys Arg Phe Ser Thr			
15 20 25			
cag aca gtt tct gaa aag cag aaa ccc ccg tcc tgg gtg atg gta acg 329			
Gln Thr Val Ser Glu Lys Gln Lys Pro Pro Ser Trp Val Met Val Thr			
30 35 40			
gat cag gaa aag cat gga aca ctt ccg gtc tac gaa gat aac atg cca 377			
Asp Gln Glu Lys His Gly Thr Leu Pro Val Tyr Glu Asp Asn Met Pro			
45 50 55			
aca ttc aac gga aaa cac cca ttg gtg aaa aca gat tca att atc ctg 425			
Thr Phe Asn Gly Lys His Pro Leu Val Lys Thr Asp Ser Ile Ile Leu			
60 65 70 75			
aaa tgt ctt ctg tcg gcc tgc ctt gtt ctc gtt tca gct ata gcc tat 473			
Lys Cys Leu Leu Ser Ala Cys Leu Val Leu Val Ser Ala Ile Ala Tyr			
80 85 90			
aaa aca aac att gga ccc gtc agt cag att aaa ccc gcc gta gcc aaa 521			
Lys Thr Asn Ile Gly Pro Val Ser Gln Ile Lys Pro Ala Val Ala Lys			
95 100 105			
acc ttt gaa act gaa ttt caa ttt gct tca gca agc cat tgg ttc gaa 569			
Thr Phe Glu Thr Glu Phe Gln Phe Ala Ser Ala Ser His Trp Phe Glu			
110 115 120			

acc aaa ttc gga aat ccg ctt gct ttc ctg gct cct gaa cac aaa aat	617
Thr Lys Phe Gly Asn Pro Leu Ala Phe Leu Ala Pro Glu His Lys Asn	
125 130 135	
aag gaa cag cag att gaa gta ggc aaa gat ctg atc gcg cct gca tcc	665
Lys Glu Gln Gln Ile Glu Val Gly Lys Asp Leu Ile Ala Pro Ala Ser	
140 145 150 155	
ggg aaa gta cag cag gat ttt cag gac aat ggg gaa gga att aaa gtc	713
Gly Lys Val Gln Gln Asp Phe Gln Asp Asn Gly Glu Gly Ile Lys Val	
160 165 170	
gaa aca agc agt gat aag att gat agc gta aaa gaa ggc tat gtg gtt	761
Glu Thr Ser Ser Asp Lys Ile Asp Ser Val Lys Glu Gly Tyr Val Val	
175 180 185	
gaa gtc agc aaa gac agc caa acg gga ctg acg gtt aag gtg cag cat	809
Glu Val Ser Lys Asp Ser Gln Thr Gly Leu Thr Val Lys Val Gln His	
190 195 200	
gct gac aac acc tat agt atc tat ggc gag ctc aaa gat gtg gat gtt	857
Ala Asp Asn Thr Tyr Ser Ile Tyr Gly Glu Leu Lys Asp Val Asp Val	
205 210 215	
gct tta tat gat ttt gtg gat aaa ggc aaa aag ctc ggt tcg att aag	905
Ala Leu Tyr Asp Phe Val Asp Lys Gly Lys Leu Gly Ser Ile Lys	
220 225 230 235	
ctt gat gat cat aat aaa ggg gtc tat tat ttt gcc atg aaa gac ggc	953
Leu Asp Asp His Asn Lys Gly Val Tyr Tyr Phe Ala Met Lys Asp Gly	
240 245 250	
gat aaa ttt att gat ccg att cag gtg att tca ttt gaa taa	995
Asp Lys Phe Ile Asp Pro Ile Gln Val Ile Ser Phe Glu	
255 260	
atggctcgac cttatcttaa agatccatgt gcattttttt ctttggattt ttgcggcgct	1055
gggcttgctc acaggccata tgaaagcatt attatgtctg ctcctgattt tattgattca	1115
tgagctgggg catgctgctc tggctgtgtt tttttcttgg agaatcaagc gtgtttttt	1175
gctgccgttt ggcggaa	1192

<210> 16
 <211> 264
 <212> PRT
 <213> *Bacillus subtilis*

<400> 16

Met Ser His Arg Ala Asp Glu Ile Arg Lys Arg Leu Glu Lys Arg Arg	
1 5 10 15	

Lys Gln Leu Ser Gly Ser Lys Arg Phe Ser Thr Gln Thr Val Ser Glu	
20 25 30	

Lys Gln Lys Pro Pro Ser Trp Val Met Val Thr Asp Gln Glu Lys His
35 40 45

Gly Thr Leu Pro Val Tyr Glu Asp Asn Met Pro Thr Phe Asn Gly Lys
50 55 60

His Pro Leu Val Lys Thr Asp Ser Ile Ile Leu Lys Cys Leu Leu Ser
65 70 75 80

Ala Cys Leu Val Leu Val Ser Ala Ile Ala Tyr Lys Thr Asn Ile Gly
85 90 95

Pro Val Ser Gln Ile Lys Pro Ala Val Ala Lys Thr Phe Glu Thr Glu
100 105 110

Phe Gln Phe Ala Ser Ala Ser His Trp Phe Glu Thr Lys Phe Gly Asn
115 120 125

Pro Leu Ala Phe Leu Ala Pro Glu His Lys Asn Lys Glu Gln Gln Ile
130 135 140

Glu Val Gly Lys Asp Leu Ile Ala Pro Ala Ser Gly Lys Val Gln Gln
145 150 155 160

Asp Phe Gln Asp Asn Gly Glu Gly Ile Lys Val Glu Thr Ser Ser Asp
165 170 175

Lys Ile Asp Ser Val Lys Glu Gly Tyr Val Val Glu Val Ser Lys Asp
180 185 190

Ser Gln Thr Gly Leu Thr Val Lys Val Gln His Ala Asp Asn Thr Tyr
195 200 205

Ser Ile Tyr Gly Glu Leu Lys Asp Val Asp Val Ala Leu Tyr Asp Phe
210 215 220

Val Asp Lys Gly Lys Lys Leu Gly Ser Ile Lys Leu Asp Asp His Asn
225 230 235 240

Lys Gly Val Tyr Tyr Phe Ala Met Lys Asp Gly Asp Lys Phe Ile Asp
245 250 255

Pro Ile Gln Val Ile Ser Phe Glu
260

<210> 17
 <211> 1264
 <212> DNA
 <213> *Bacillus subtilis*

<220>
 <221> CDS
 <222> (201)..(1067)
 <223>

<220>
 <221> gene
 <222> (1)..(1264)
 <223> spoIVFB

<220>
 <221> misc_feature
 <222> (201)..(203)
 <223> First codon translated as Met.

<400> 17
 actgacgggt aaggtgcagc atgctgacaa cacctatagt atctatggcg agctcaaaga 60
 tgtggatgtt gctttatatg attttgtgga taaaggcaaa aagctcggtt cgattaagct
 tgatgatcat aataaagggg tctattattt tgccatgaaa gacggcgata aatttattga 120
 tccgattcag gtgatttcat ttg aat aaa tgg ctc gac ctt atc tta aag atc
 Leu Asn Lys Trp Leu Asp Leu Ile Leu Lys Ile 233
 1 5 10

cat gtg cat cct ttt ctt tgg att att gcg gcg ctg ggc ttg ctc aca 281
 His Val His Pro Phe Leu Trp Ile Ile Ala Ala Leu Gly Leu Leu Thr
 15 20 25

ggc cat atg aaa gca tta tta tgt ctg ctc ctg att gta ttg att cat 329
 Gly His Met Lys Ala Leu Leu Cys Leu Leu Ile Val Leu Ile His
 30 35 40

gag ctg ggg cat gct gct ctg gct gtg ttt ttt tct tgg aga atc aag 377
 Glu Leu Gly His Ala Ala Leu Ala Val Phe Phe Ser Trp Arg Ile Lys
 45 50 55

cgt gtt ttt ttg ctg ccg ttt ggc gga acg gtc gaa gtg gaa gag cac 425
 Arg Val Phe Leu Leu Pro Phe Gly Gly Thr Val Glu Val Glu Glu His
 60 65 70 75

ggg aat cgg ccg tta aag gaa gag ttt gcg gtc att att gcc gga cct 473
 Gly Asn Arg Pro Leu Lys Glu Glu Phe Ala Val Ile Ile Ala Gly Pro
 80 85 90

ctt cag cac atc tgg ctt cag ttt gcc gcc tgg atg ctt gca gaa gtc 521
 Leu Gln His Ile Trp Leu Gln Phe Ala Ala Trp Met Leu Ala Glu Val
 95 100 105

tca	gtg	att	cat	cag	cat	acc	ttt	gaa	ctc	ttc	acc	ttt	tat	aat	ctt	569
Ser	Val	Ile	His	Gln	His	Thr	Phe	Glu	Leu	Phe	Thr	Phe	Tyr	Asn	Leu	
110							115						120			
tct	atc	tta	ttt	gtc	aat	tta	ctg	ccg	atc	tgg	ccg	ctg	gat	gga	gga	617
Ser	Ile	Leu	Phe	Val	Asn	Leu	Leu	Pro	Ile	Trp	Pro	Leu	Asp	Gly	Gly	
125							130						135			
aaa	ctg	tta	ttt	ttg	ttg	ttt	tcc	aaa	cag	ctg	cct	ttt	caa	aag	gct	665
Lys	Leu	Leu	Phe	Leu	Leu	Phe	Ser	Lys	Gln	Leu	Pro	Phe	Gln	Lys	Ala	
140							145					150		155		
cac	ccg	ctt	aat	cta	aaa	acg	tcg	ctc	tgc	ttc	tgc	ctg	ctg	ctc	ggg	713
His	Arg	Ileu	Asn	Leu	Lys	Thr	Ser	Leu	Cys	Phe	Cys	Leu	Leu	Gly		
160							165						170			
tgc	tgg	gtt	tta	ttc	gtg	att	cct	ctg	caa	atc	agc	gca	tgg	gtt	ttg	761
Cys	Trp	Val	Leu	Phe	Val	Ile	Pro	Leu	Gln	Ile	Ser	Ala	Trp	Val	Leu	
175							180						185			
ttt	gtc	ttt	ctg	gct	gtt	tcc	ttg	ttt	gag	gaa	tat	cg	caa	agg	cac	809
Phe	Val	Phe	Leu	Ala	Val	Ser	Leu	Phe	Glu	Glu	Tyr	Arg	Gln	Arg	His	
190							195					200				
tat	atc	cat	gtg	aga	ttt	ctc	ctc	gaa	agg	tat	tac	gga	aaa	aac	agg	857
Tyr	Ile	His	Val	Arg	Phe	Leu	Leu	Glu	Arg	Tyr	Tyr	Gly	Lys	Asn	Arg	
205							210					215				
gag	ctt	gag	aaa	ctt	ctg	ccg	ctg	aca	gta	aag	g	g	g	g	aaa	905
Glu	Leu	Glu	Leu	Leu	Leu	Pro	Leu	Thr	Val	Lys	Ala	Glu	Asp	Lys	Val	
220							225					230		235		
tat	cat	gtg	atg	gcc	gag	ttc	aaa	cgt	ggc	tgt	aag	cat	ccg	att	att	953
Tyr	His	Val	Met	Ala	Glu	Phe	Lys	Arg	Gly	Cys	Lys	His	Pro	Ile	Ile	
240							245						250			
ata	gaa	aaa	tca	ggc	caa	aag	ctc	agc	cag	ctt	gac	gag	aat	gaa	gtg	1001
Ile	Glu	Lys	Ser	Gly	Gln	Lys	Leu	Ser	Gln	Leu	Asp	Glu	Asn	Glu	Val	
255							260						265			
ctg	cac	gct	tac	ttt	gcc	gat	aag	cg	ac	aat	tct	tcc	atg	gag	gaa	1049
Leu	His	Ala	Tyr	Phe	Ala	Asp	Lys	Arg	Thr	Asn	Ser	Ser	Met	Glu	Glu	
270							275						280			
ctg	ctt	ctg	ccc	tac	taa	aactgattga	caa	acgc	ttt	gtat	ttt	gg	ttt	gg	ttt	1097
Leu	Leu	Leu	Pro	Tyr												
285																
atat	tttta	atgttatgga	tgt	tagc	acca	tt	gct	aca	ac	cg	tc	ag	tg	tt	aa	1157
ag	ctt	tttaca	cccc	cct	tgt	at	ct	ggc	gag	t	ctt	ag	tg	tt	aa	1217
ac	gca	aat	cat	taaa	acag	gc	gg	taaa	aca	aa	tca	aa	gt	tt	ga	1264

<210> 18
<211> 288
<212> PRT

<213> Bacillus subtilis

<220>

<221> misc_feature

<222> (201)..(203)

<223> First codon translated as Met.

<400> 18

Leu Asn Lys Trp Leu Asp Leu Ile Leu Lys Ile His Val His Pro Phe
1 5 10 15

Leu Trp Ile Ile Ala Ala Leu Gly Leu Leu Thr Gly His Met Lys Ala
20 25 30

Leu Leu Cys Leu Leu Leu Ile Val Leu Ile His Glu Leu Gly His Ala
35 40 45

Ala Leu Ala Val Phe Phe Ser Trp Arg Ile Lys Arg Val Phe Leu Leu
50 55 60

Pro Phe Gly Gly Thr Val Glu Val Glu His Gly Asn Arg Pro Leu
65 70 75 80

Lys Glu Glu Phe Ala Val Ile Ile Ala Gly Pro Leu Gln His Ile Trp
85 90 95

Leu Gln Phe Ala Ala Trp Met Leu Ala Glu Val Ser Val Ile His Gln
100 105 110

His Thr Phe Glu Leu Phe Thr Phe Tyr Asn Leu Ser Ile Leu Phe Val
115 120 125

Asn Leu Leu Pro Ile Trp Pro Leu Asp Gly Gly Lys Leu Leu Phe Leu
130 135 140

Leu Phe Ser Lys Gln Leu Pro Phe Gln Lys Ala His Arg Leu Asn Leu
145 150 155 160

Lys Thr Ser Leu Cys Phe Cys Leu Leu Leu Gly Cys Trp Val Leu Phe
165 170 175

Val Ile Pro Leu Gln Ile Ser Ala Trp Val Leu Phe Val Phe Leu Ala
180 185 190

Val Ser Leu Phe Glu Glu Tyr Arg Gln Arg His Tyr Ile His Val Arg
195 200 205

Phe Leu Leu Glu Arg Tyr Tyr Gly Lys Asn Arg Glu Leu Glu Lys Leu
210 215 220

Leu Pro Leu Thr Val Lys Ala Glu Asp Lys Val Tyr His Val Met Ala
225 230 235 240

Glu Phe Lys Arg Gly Cys Lys His Pro Ile Ile Ile Glu Lys Ser Gly
245 250 255

Gln Lys Leu Ser Gln Leu Asp Glu Asn Glu Val Leu His Ala Tyr Phe
260 265 270

Ala Asp Lys Arg Thr Asn Ser Ser Met Glu Glu Leu Leu Leu Pro Tyr
275 280 285

<210> 19
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> PCR-Primer spol

<400> 19
ggctgatgct caaacagggg cagtgcatc

29

<210> 20
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer spo2

<400> 20
catgaacggc ctttacgaca gcc

24

<210> 21
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer spo3

<400> 21
gtcatcaaaa cgattttgcc tgagg

25

<210> 22

<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer spo4

<400> 22
atgttctgtc ccgggattgg ctccctg

26

<210> 23
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer spo6

<400> 23
gttttgactc tgatcggaat tctttggcg

29

<210> 24
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer spo7

<400> 24
gcacgaaacg agcgagaatg gc

22

<210> 25
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer recA1

<400> 25
ggaattcggc atcagcttca ctggag

26

<210> 26
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer recA2

<400> 26
gctatgtcga ctataccctg tttatgcgg

29

<210> 27
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer recA3

<400> 27
gacctcgaa cagagcttga c

21

<210> 28
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer recA4

<400> 28
tcaaactgca gtcattaaga gaatggatgg

30

<210> 29
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer recA5

<400> 29
aagcttacgg tttaacgttt ctg

23

<210> 30
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR-Primer recA6

<400> 30
acacaaacga attgaaagtgcagcg

26

<210> 31
<211> 1557
<212> DNA
<213> *Bacillus licheniformis* A

<220>
<221> CDS
<222> (369)..(1415)
<223>

```

<220>
<221> gene
<222> (1)..(1557)
<223> recA

<400> 31
gatatcgcca tcagttcac tggagtagcc gggccgaata cgcaagaagg ccatccggct 60
ggaaagggtgt ttatcgcat ctccgtgaag gaccaggctg aggaagcggt cgaatttcag 120
ttcgccggat ccaggtctgc ggtgcggaaag cgttctgcca aatacggctg ccatctgttg 180
ctgaaaatga tggaaaaata agcggaaacc ggattttcgg aatatcttc tttcgaaaaa 240
ggccattcca ttttaggaga tcgattttc ctctaaaaaa aatcgaatat gcgttcgctt 300
tttcttggc aaatccgcat aaacaaggta tagtagatat agcggaaagtg ataaaggagg 360
aaaataga atg agt gat cgt cag gca gcc tta gat atg gcg ctt aaa caa 410
Met Ser Asp Arg Gln Ala Ala Leu Asp Met Ala Leu Lys Gln
1 5 10

ata gaa aag cag ttt ggt aaa ggt tcg att atg aaa ctc ggc gaa caa 458
Ile Glu Lys Gln Phe Gly Lys Ser Ile Met Lys Leu Gly Glu Gln
15 20 25 30

act gaa acg aga att tca aca gtt ccg agc ggt tct tta gcg ctc gat 506
Thr Glu Thr Arg Ile Ser Thr Val Pro Ser Gly Ser Leu Ala Leu Asp
35 40 45

gcg gct ctt gga gtg ggc gga tac ccg cgc ggc cgg att att gaa gta 554
Ala Ala Leu Gly Val Gly Gly Tyr Pro Arg Gly Arg Ile Ile Glu Val
50 55 60

tac ggg cct gaa agc tcc ggt aaa acg acg gtg gcg ctt cat gcg att 602
Tyr Gly Pro Glu Ser Ser Gly Lys Thr Thr Val Ala Leu His Ala Ile
65 70 75

gcc gaa gtt cag cag cag ggc gga caa gcg gcg ttc atc gac gcc gaa 650
Ala Glu Val Gln Gln Gly Gly Gln Ala Ala Phe Ile Asp Ala Glu
80 85 90

cac gcg ctt gat ccc gtc tat gca caa aag ctg ggc gtc aac att gat 698
His Ala Leu Asp Pro Val Tyr Ala Gln Lys Leu Gly Val Asn Ile Asp
95 100 105 110

gag ctt ttg ctg tca cag cct gat acg ggc gag cag gcg ctc gaa atc 746
Glu Leu Leu Ser Gln Pro Asp Thr Gly Glu Gln Ala Leu Glu Ile
115 120 125

gct gaa gcc ctt gtc aga agc gga gcg gtg gat atc gtt gtc atc gac 794
Ala Glu Ala Leu Val Arg Ser Gly Ala Val Asp Ile Val Val Ile Asp
130 135 140

tct gta gca gcg ctt gtg ccg aaa gct gaa atc gaa gga gat atg ggg 842
Ser Val Ala Ala Leu Val Pro Lys Ala Glu Ile Glu Gly Asp Met Gly
145 150 155

```

gat tcc cac gtc ggt ttg cag gcc aga ctg atg tct cag gcg ctt cgc	890
Asp Ser His Val Gly Leu Gln Ala Arg Leu Met Ser Gln Ala Leu Arg	
160 165 170	
aag ctt tcc gga gcg atc aat aaa tcg aag acc atc gcg atc ttt atc	938
Lys Leu Ser Gly Ala Ile Asn Lys Ser Lys Thr Ile Ala Ile Phe Ile	
175 180 185 190	
aac cag att cgt gaa aaa gtc ggt gtc atg ttt gga aat cct gag acg	986
Asn Gln Ile Arg Glu Lys Val Gly Val Met Phe Gly Asn Pro Glu Thr	
195 200 205	
acg cca ggc gga aga gcg ctg aaa ttc tac tct tct gtc cgc ctt gaa	1034
Thr Pro Gly Gly Arg Ala Leu Lys Phe Tyr Ser Ser Val Arg Leu Glu	
210 215 220	
gtg cgc cgc gca gag cag ctg aaa caa ggc aac gac gtc atg ggg aac	1082
Val Arg Arg Ala Glu Gln Leu Lys Gln Gly Asn Asp Val Met Gly Asn	
225 230 235	
aag acg aaa atc aaa gtc gtg aaa aac aaa gtg gca cct cca ttc cgg	1130
Lys Thr Lys Ile Lys Val Val Lys Asn Lys Val Ala Pro Pro Phe Arg	
240 245 250	
aca gcc gaa gtg gac att atg tac ggg gaa gga att tca aaa gaa ggg	1178
Thr Ala Glu Val Asp Ile Met Tyr Gly Glu Gly Ile Ser Lys Glu Gly	
255 260 265 270	
gaa atc atc gac ctc gga aca gag ctt gac atc gtt caa aag agc ggt	1226
Glu Ile Ile Asp Leu Gly Thr Glu Leu Asp Ile Val Gln Lys Ser Gly	
275 280 285	
gca tgg tac tct tat cag gag gaa cgc ctt gga caa ggc cgt gaa aac	1274
Ala Trp Tyr Ser Tyr Gln Glu Glu Arg Leu Gly Gln Gly Arg Glu Asn	
290 295 300	
gcc aaa cag ttc ctg aaa gaa aac aag gat atc ctt ttg atg att caa	1322
Ala Lys Gln Phe Leu Lys Glu Asn Lys Asp Ile Leu Leu Met Ile Gln	
305 310 315	
gag cag atc cgg gag cac tac ggt ttg gat act gga ggc gct gct cct	1370
Glu Gln Ile Arg Glu His Tyr Gly Leu Asp Thr Gly Gly Ala Ala Pro	
320 325 330	
gca cag gaa gac gag gcc caa gct cag gaa gaa ctc gag ttt taa	1415
Ala Gln Glu Asp Glu Ala Gln Ala Gln Glu Glu Leu Glu Phe	
335 340 345	
tcatgaaacg tgtgaaaggc tgccggcccg atcggcagcc ttttacttta ttcttcgctt	1475
tcaggcgctt ctcttccatc cattctctta atgagggcag tttgaaaggc gtttaatcca	1535
gaaacgttaa gaccgtaagc tt	1557

<210> 32

<211> 348

<212> PRT

<213> *Bacillus licheniformis* A

<400> 32

Met Ser Asp Arg Gln Ala Ala Leu Asp Met Ala Leu Lys Gln Ile Glu
1 5 10 15

Lys Gln Phe Gly Lys Gly Ser Ile Met Lys Leu Gly Glu Gln Thr Glu
20 25 30

Thr Arg Ile Ser Thr Val Pro Ser Gly Ser Leu Ala Leu Asp Ala Ala
35 40 45

Leu Gly Val Gly Gly Tyr Pro Arg Gly Arg Ile Ile Glu Val Tyr Gly
50 55 60

Pro Glu Ser Ser Gly Lys Thr Thr Val Ala Leu His Ala Ile Ala Glu
65 70 75 80

Val Gln Gln Gln Gly Gly Gln Ala Ala Phe Ile Asp Ala Glu His Ala
85 90 95

Leu Asp Pro Val Tyr Ala Gln Lys Leu Gly Val Asn Ile Asp Glu Leu
100 105 110

Leu Leu Ser Gln Pro Asp Thr Gly Glu Gln Ala Leu Glu Ile Ala Glu
115 120 125

Ala Leu Val Arg Ser Gly Ala Val Asp Ile Val Val Ile Asp Ser Val
130 135 140

Ala Ala Leu Val Pro Lys Ala Glu Ile Glu Gly Asp Met Gly Asp Ser
145 150 155 160

His Val Gly Leu Gln Ala Arg Leu Met Ser Gln Ala Leu Arg Lys Leu
165 170 175

Ser Gly Ala Ile Asn Lys Ser Lys Thr Ile Ala Ile Phe Ile Asn Gln
180 185 190

Ile Arg Glu Lys Val Gly Val Met Phe Gly Asn Pro Glu Thr Thr Pro
195 200 205

Gly Gly Arg Ala Leu Lys Phe Tyr Ser Ser Val Arg Leu Glu Val Arg
210 215 220

Arg Ala Glu Gln Leu Lys Gln Gly Asn Asp Val Met Gly Asn Lys Thr
225 230 235 240

Lys Ile Lys Val Val Lys Asn Lys Val Ala Pro Pro Phe Arg Thr Ala
245 250 255

Glu Val Asp Ile Met Tyr Gly Glu Gly Ile Ser Lys Glu Gly Glu Ile
260 265 270

Ile Asp Leu Gly Thr Glu Leu Asp Ile Val Gln Lys Ser Gly Ala Trp
275 280 285

Tyr Ser Tyr Gln Glu Glu Arg Leu Gly Gln Gly Arg Glu Asn Ala Lys
290 295 300

Gln Phe Leu Lys Glu Asn Lys Asp Ile Leu Leu Met Ile Gln Glu Gln
305 310 315 320

Ile Arg Glu His Tyr Gly Leu Asp Thr Gly Gly Ala Ala Pro Ala Gln
325 330 335

Glu Asp Glu Ala Gln Ala Gln Glu Glu Leu Glu Phe
340 345

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2005/001543

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/195

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, MEDLINE, WPI Data, PAJ, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 02/29113 A (NOVOZYMES BIOTECH, INC; NOVOZYMES A/S) 11 April 2002 (2002-04-11)	38, 43
X	page 164, line 16	38
X	page 174, line 16	43
X	----- KUNST F ET AL: "THE COMPLETE GENOME SEQUENCE OF THE GRAM-POSITIVE BACTERIUM BACILLUS SUBTILIS" NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, vol. 390, 1997, pages 249-256, XP000919353 ISSN: 0028-0836 the whole document ----- -/-	1-5, 38, 43

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

° Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
2 June 2005	13/06/2005
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2260 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Stoyanov, B

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2005/001543

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	VEITH B ET AL.: "The complete genome sequence of <i>Bacillus Licheniformis</i> DSM13, an organism with great industrial potential" JOURNAL OF MOLECULAR MICROBIOLOGY AND BIOTECHNOLOGY, vol. 7, no. 4, 31 March 2004 (2004-03-31), pages 204-211, XP009047713 the whole document -----	1-7, 31, 32

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP2005/001543

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 0229113	A 11-04-2002	AU 9671801	A 15-04-2002	EP 1355931 A2 29-10-2003
		WO 0229113	A2 11-04-2002	US 2002146721 A1 10-10-2002

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

CT/EP2005/001543

A. KLASIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 7 C07K14/195

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchiertes Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
IPK 7 C07K

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, MEDLINE, WPI Data, PAJ, BIOSIS

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	WO 02/29113 A (NOVOZYMES BIOTECH, INC; NOVOZYMES A/S) 11. April 2002 (2002-04-11)	38,43
X	Seite 164, Zeile 16	38
X	Seite 174, Zeile 16	43
X	KUNST F ET AL: "THE COMPLETE GENOME SEQUENCE OF THE GRAM-POSITIVE BACTERIUM BACILLUS SUBTILIS" NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, Bd. 390, 1997, Seiten 249-256, XP000919353 ISSN: 0028-0836 das ganze Dokument	1-5,38, 43
	----- -/-	

Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

Siehe Anhang Patentfamilie

- ° Besondere Kategorien von angegebenen Veröffentlichungen :
- **A** Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist
- **E** älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist
- **L** Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)
- **O** Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht
- **P** Veröffentlichung, die vor dem internationalem Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist
- **T** Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist
- **X** Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erforderliche Tätigkeit beruhend betrachtet werden
- **Y** Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erforderlicher Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann nahelegend ist
- **Z** Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche	Absendedatum des internationalen Recherchenberichts
2. Juni 2005	13/06/2005
Name und Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bevollmächtigter Bediensteter Stoyanov, B

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2005/001543

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie ^a	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
P, X	VEITH B ET AL.: "The complete genome sequence od <i>Bacillus Licheniformis</i> DSM13, an organism with great industrial potential" JOURNAL OF MOLECULAR MICROBIOLOGY AND BIOTECHNOLOGY, Bd. 7, Nr. 4, 31. März 2004 (2004-03-31), Seiten 204-211, XP009047713 das ganze Dokument -----	1-7, 31, 32

INTERNATIONALER RECHERCHENBERICHTInternationales Aktenzeichen
PCT/EP2005/001543**Feld II Bemerkungen zu den Ansprüchen, die sich als nicht recherchierbar erwiesen haben (Fortsetzung von Punkt 2 auf Blatt 1)**

Gemäß Artikel 17(2)a) wurde aus folgenden Gründen für bestimmte Ansprüche kein Recherchenbericht erstellt:

1. Ansprüche Nr.
weil sie sich auf Gegenstände beziehen, zu deren Recherche die Behörde nicht verpflichtet ist, nämlich

2. Ansprüche Nr.
weil sie sich auf Teile der internationalen Anmeldung beziehen, die den vorgeschriebenen Anforderungen so wenig entsprechen, daß eine sinnvolle internationale Recherche nicht durchgeführt werden kann, nämlich

3. Ansprüche Nr.
weil es sich dabei um abhängige Ansprüche handelt, die nicht entsprechend Satz 2 und 3 der Regel 6.4 a) abgefaßt sind.

Feld III Bemerkungen bei mangelnder Einheitlichkeit der Erfindung (Fortsetzung von Punkt 3 auf Blatt 1)

Die internationale Recherchenbehörde hat festgestellt, daß diese internationale Anmeldung mehrere Erfindungen enthält:

1. Da der Anmelder alle erforderlichen zusätzlichen Recherchengebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht auf alle recherchierbaren Ansprüche.

2. Da für alle recherchierbaren Ansprüche die Recherche ohne einen Arbeitsaufwand durchgeführt werden konnte, der eine zusätzliche Recherchengebühr gerechtfertigt hätte, hat die Behörde nicht zur Zahlung einer solchen Gebühr aufgefordert.

3. Da der Anmelder nur einige der erforderlichen zusätzlichen Recherchengebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht nur auf die Ansprüche, für die Gebühren entrichtet worden sind, nämlich auf die Ansprüche Nr.

4. Der Anmelder hat die erforderlichen zusätzlichen Recherchengebühren nicht rechtzeitig entrichtet. Der internationale Recherchenbericht beschränkt sich daher auf die in den Ansprüchen zuerst erwähnte Erfindung; diese ist in folgenden Ansprüchen erfaßt:

Bemerkungen hinsichtlich eines Widerspruchs

Die zusätzlichen Gebühren wurden vom Anmelder unter Widerspruch gezahlt.
 Die Zahlung zusätzlicher Recherchengebühren erfolgte ohne Widerspruch.

WEITERE ANGABEN

PCT/ISA/ 210

Die internationale Recherchenbehörde hat festgestellt, dass diese internationale Anmeldung mehrere (Gruppen von) Erfindungen enthält, nämlich:

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen
PCT/EP2005/001543

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie		Datum der Veröffentlichung
WO 0229113	A 11-04-2002	AU	9671801 A	15-04-2002
		EP	1355931 A2	29-10-2003
		WO	0229113 A2	11-04-2002
		US	2002146721 A1	10-10-2002